The DNA Saga

50 Years of the Double Helix

Biman Basu

Vigyan Prasar

Published by
Vigyan Prasar
C-24, Qutab Institutional Area
New Delhi - 110 016
(Regd. office : Technology Bhawan, New Delhi - 110 016)
Phones : 26864157, 26864022 Fax : 26965986
E-mail : vigyan@hub.nic.in
Internet : http://www.vigyanprasar.com

The DNA Saga
50 Years of the Double Helix

Author : Biman Basu

Overall Supervision : *Dr. Subodh Mahanti*
Typesetting & pagemaking : *Ms. Sonu*
Production Supervision : *Sumita Sen*

ISBN- 81-7480-099-9

Rs.75.00

Printed in India by : Saurabh Print-O-Pack

Contents

Foreword

The Publication Programme of Vigyan Prasar has taken some shape in the last few years. To start with, Vigyan Prasar brought out a number of publications on a variety of topics of science and technology on an experimental basis. Popular Science classics, India's Scientific Heritage, Natural History, Health, and Do-It-Yourself are some of the series that evolved over the years. Our emphasis has been on bringing out quality publications on various aspects of science and technology at affordable prices. Further, Vigyan Prasar is putting in efforts to bring out publications in major Indian languages for various target groups.

Exacly fifty years ago, James D. Watson and Francis H. C. Crick published their monumental paper on the double helical structure of DNA. The present book, *The DNA Saga: 50 Years of Double Helix* commemorates this historic event. The author, Shri Biman Basu, is a well-known science populariser. Indeed, there is no gainsaying the fact that such a popular account is not readily available. It is hoped that even those who do not have any requisite background on the subject will enjoy the book. Suggestions for further improvement of the book will be appreciated.

V.B. Kamble
Acting Director
Vigyan Prasar

New Delhi
April 24, 2003

Preface

The discovery of the double helix structure of DNA was a turning point in the history of biology. Not only did it help us understand the very nature of life but it also opened up the possibility of controlling life artificially. Strictly speaking, however, it was only one, may be the most significant one, in a series of discoveries made through innovative experiments over several years, by dedicated teams of scientists that ultimately led to our understanding of the inner mechanisms of the life process. As fifty years of explosive study has shown, our understanding of DNA changed everything. Within a decade, it explained exactly how genes store genetic information in a four-letter chemical alphabet, how they replicate, and how they drive not only the development and functioning of the cell, but heredity and evolution as well.

As new insights were gained, new possibilities opened up, finally leading to recombinant DNA technology, or genetic engineering that is revolutionising agriculture, health care and industrial production of chemicals and pharmaceuticals. Genetically modified crops have been developed that resist pests and give higher yields. Transgenic rice, protein-enriched potato, and oxalate-free tomato have longer shelf life and provide better nutrition. Recombinant DNA based diagnostic kits give reliable diagnoses must faster than conventional methods. The complete mapping of the human genome has opened up enormous possibilities of diagnosing and treating intractable diseases like multiple sclerosis and cancer. And the list is still growing. This book is an attempt at presenting the exciting story of events leading to the discovery of DNA structure and the subsequent developments in molecular biology that have touched every aspect of our life.

Biman Basu

New Delhi
March 15, 2003

Acknowledgements

The idea of bringing out this book to mark the golden jubilee of Watson and Crick's momentous discovery was mooted by my long-time friend and at present Acting Director of Vigyan Prasar, Dr. Vinay B. Kamble. When he asked me to write it, I was a little hesitant, as the time he gave was rather short. But I agreed, and I must say, without his encouragement and prodding I would never have ventured to go ahead. Now that the job is done, I feel happy to have been able to keep to the deadline. I am thankful to Dr. Kamble for giving me this opportunity. I am fortunate to have the manuscript perused by Dr. Subodh Mahanti of Vigyan Prasar, who himself has worked on Molecular Biology. I am grateful to him for taking the trouble of going through the manuscript and offering valuable tips. My thanks are also due to Ms. Sonu and Ms. Sumita Sen of Vigyan Prasar, without whose cooperation it would have been impossible to bring out the book in time. I am also grateful to many others who have offered useful suggestions that have helped me bring the book to its present shape. My sincere thanks to all of them.

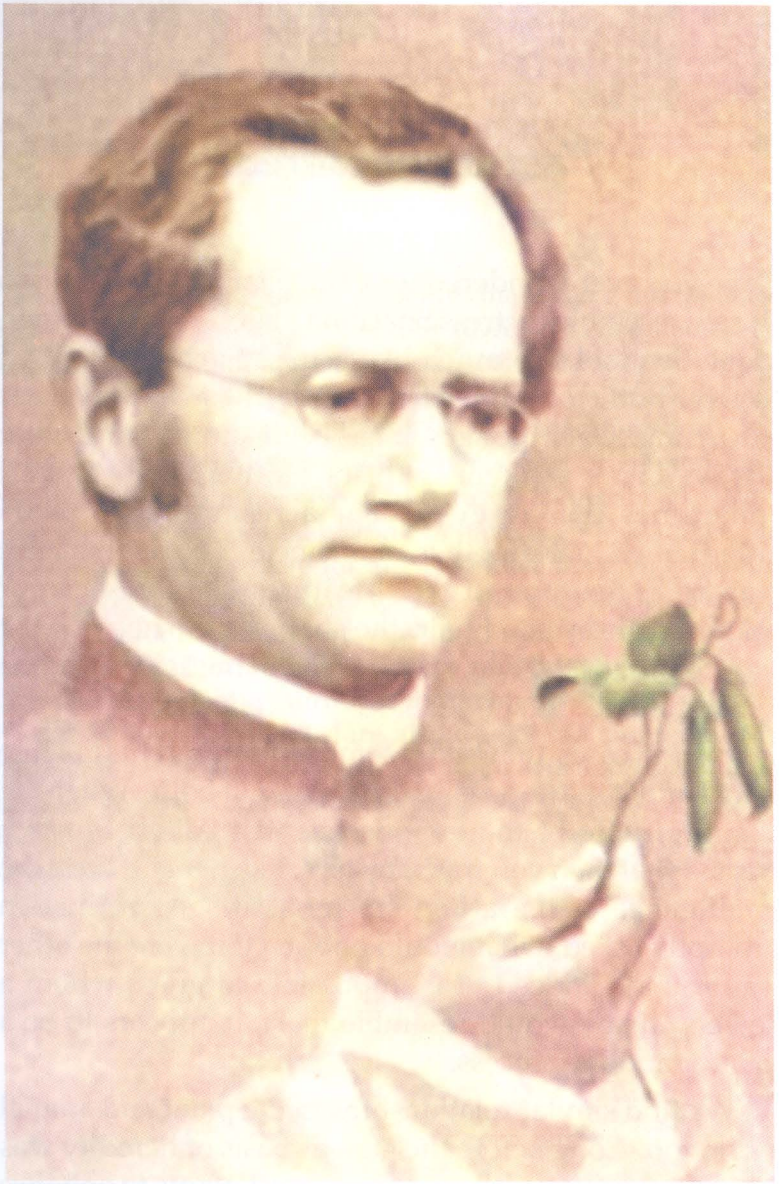
1

Factors of Heredity

Have you ever wondered why film star Saif Ali Khan so closely resembles actress Sharmila Tagore, or why Moon Moon Sen looks so much like actress Suchitra Sen? Of course, both look like their mothers. And that's not surprising. Over the ages, it has been known that a child always bears a likeness to its mother or its father; sometimes it may even bear a resemblance to its grandfather, grandmother, or an uncle. The resemblance may be in the form of facial features like the shape or colour of the eye and hair, the shape of the nose, or the skin colour, or the hair on the scalp, sometimes even the temperament or gait.

In most cases, the child resembles both parents to some extent and it may appear to be a blend of both. When the beautiful wife of the famous scientist Albert Einstein wished that their coming baby look like her and have the intelligence of its father, so the story goes, Einstein promptly asked, what would happen if it were the other way round! Indeed, it is impossible to guess whom a newborn baby would resemble most; it appears to be a totally random process.

We find similar relationships in the plant and animal world too. Long before the genetic basis of heredity was understood, farmers bred animals and plants to enhance desirable traits. Centuries of breeding of domestic plants and animals have shown that useful traits such as yield of milk in milch cattle, yield of crops such as wheat and rice, and even running ability in race horses can be



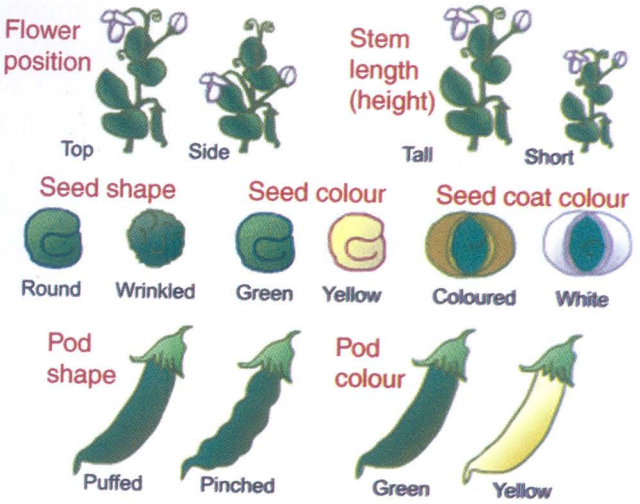
Gregor Mendel, who laid the foundation of genetics by his studies with the pea plant

transferred by selective mating. Here, too, there is no scientific way to predict the exact outcome of a cross between two particular parents. Sometimes it would work, sometimes it wouldn't. That's why it takes hundreds of trials, spread over many years, by animal and plant breeders to get a particularly desirable trait in the hybrid.

But why do the offspring always resemble their parents and carry their traits? This question had long remained an enigma, till an Augustinian monk named Gregor Mendel made a profound discovery. Working with pea plants in a small monastery garden, he discovered that individual traits are determined by discrete "factors" (later known as genes), which are inherited from the parents. He chose pea plants because they are easy to grow and they show distinct traits in plant height, pod shape, and seed shape and colour, which can be studied easily.

Mendel was no ordinary monk, nor was he an ordinary gardener. By careful observation, he was able to identify as many as seven distinct traits, namely, stem length, flower position, pod shape, pod colour, seed shape, seed colour, and seed coat colour in pea plants, which today we know as phenotypes. He focussed on the seven individual traits separately rather than looking at the pea plant as a whole, to do his crossbreeding experiment.

Mendel noticed that each of the traits visible in the pea plant had two alternate forms. For example, the plant could be "short" or "tall"; the flower position could be "top" or "side"; pod shape could be "puffed" or "pinched"; the pod colour could be "green" or "yellow"; seed shape could be "round" or "wrinkled"; seed colour could be "green" or "yellow"; and the seed coat could be "coloured" or "white". Mendel reasoned that each trait

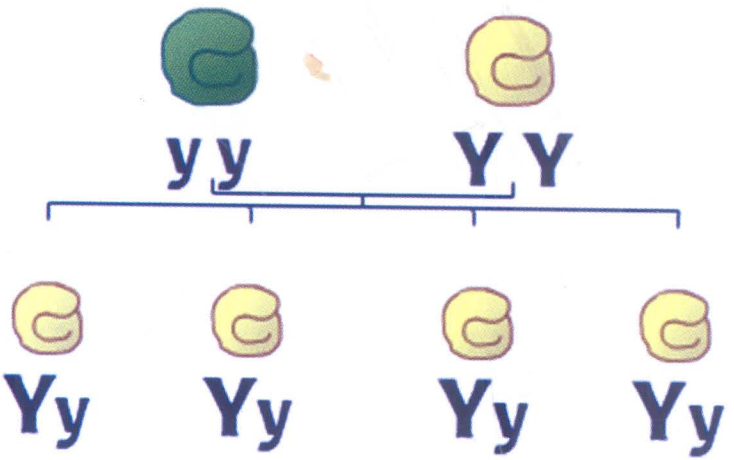


The seven traits of the pea plant

is controlled by one factor, which has two forms. We know them as alleles and the pair of alleles is called the genotype.

Using purebred strains, Mendel carefully crossbred plants with distinct traits and analysed the results. His results were startling. He found that when pea plants were crossbred by careful hand pollination of flowers, they did not produce offspring with blended traits. For example, plants bearing green seeds when crossed with plants bearing yellow seeds did not produce seeds with intermediate green-yellow colour, but produced only yellow seeds. Similarly, plants bearing puffed pods when crossed with plants with pinched pods did not produce "puffed-pinched" pods, but produced only puffed pods.

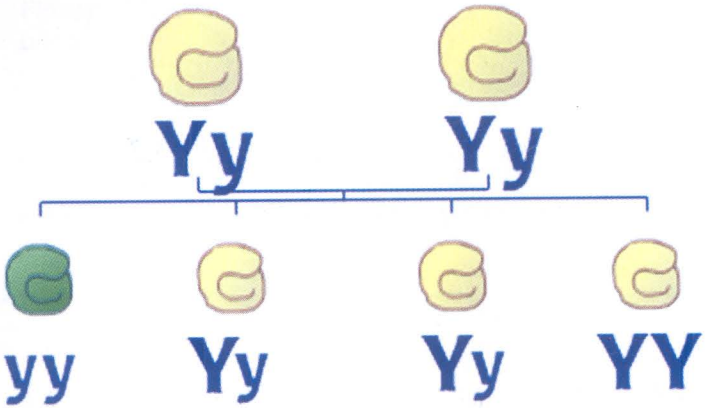
To explain the results, Mendel proposed that the "factors" responsible for different traits in pea plant come in pairs, one of which is "dominant" over the other and that each alternative form of a trait is specified by



Crossing of plants bearing green seeds with plants bearing yellow seeds produced only yellow seeds

alternative forms of a gene. In the above examples, the allele that specifies yellow colour happens to be dominant over that for green, while the allele that specifies puffed pods is dominant over that for pinched pods. So they overpower the weaker, or "recessive" traits that specify green colour in seeds and pinched pods.

Further experiments with crossbred plants brought out more surprises. When Mendel crossed the yellow seed hybrids with the same variety, some of the pods produced green seeds! How the crossing of yellow seeds could produce green was certainly puzzling. But, as a true scientist, Mendel repeated the experiment several times and finally figured out the basic rules of inheritance that could explain why the green colour appeared in the second generation. He proposed that the dominant trait is seen whenever a single copy of the factor responsible for it is inherited and that the recessive trait manifests itself only when a copy of the factor responsible for it is



Crossing of yellow seed hybrids with the same variety produced green seeds in the ratio 1:3

inherited from both the parents.

Mendel found that different combination of factors in the parents resulted in specific ratios of dominant-to-recessive traits. He could work out the results of a cross between two hybrid parents – each carrying one dominant and one recessive allele. In case of the pea plant, for example, he showed that the crossing of two yellow-seed hybrids would produce three times as many yellow seeds as green seeds. This came to be known as Mendel's famous "3 to 1" ratio.

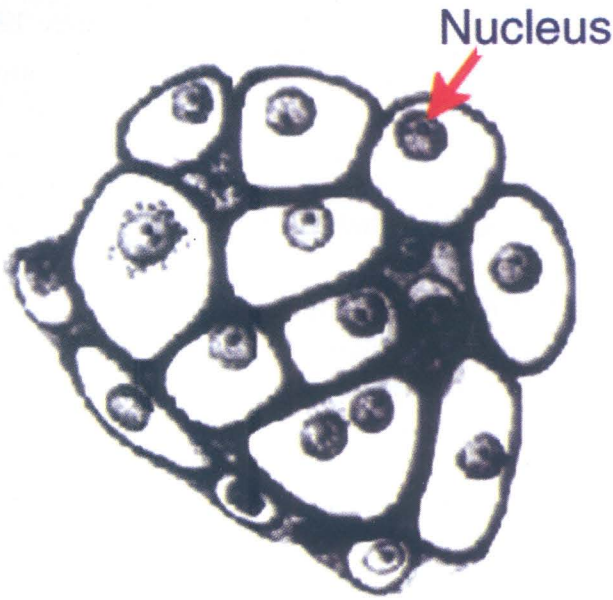
Mendel published the results of his experiments with pea plants under the title *Experiments in Plant Hybrids* in the proceedings of the Natural Science Society in Brünn (now Brno, Czech Republic) in 1865. But, as happens with many a path-breaking discovery in science, at the time his work was published no one really cared. Naturalists of his time were mostly interested in classifying living things from their outward appearance; they were not interested in finding out the cause of their appearance. Mendel's path-breaking discovery remained

mostly unknown till it was revived in the early 1900s.

The resurrection of Mendel's laws came through the work of three scientists – Dutch botanist Hugo de Vries, German botanist Carl Erich Correns and Austrian botanist Erich Tschermak von Seysenegg, each of whom worked independently and published their work simultaneously in 1900.

Hugo de Vries worked with flowers of the evening primrose and found marked differences between the wild variety and the cultivated species, which he thought were caused by mutations. His research into the nature of mutations subsequently led him to begin a program of plant breeding in 1892, and eight years later he drew up the same laws of heredity that Mendel had, 35 years earlier. Correns did his research with garden peas while an instructor of botany at the University of Tübingen. From his results he drew the same conclusions that Mendel had. Surveying the literature on the subject, he discovered Mendel's original paper. Tschermak also experimented with the garden pea, in the Botanical Garden of Ghent in the spring of 1898. While writing the results of his experiments, Tschermak saw a cross-reference to Mendel's work. He found that Mendel's work with the garden pea duplicated and in some ways superseded his own. So, by 1900, Mendel's laws of heredity were placed on a sound footing.

Meanwhile, the invention of the microscope had revealed the cellular nature of all living matter. There was strong evidence that cells were the basic units of life. As early as 1839, a German physiologist named Theodor Schwann had discovered that all living cells contained a nucleus. Later, it was found that during cell division (mitosis), the nuclei break up into small, distinctive thread-like structures. These structures were subsequently found to absorb certain dyes and so came



All living cells have a nucleus that contains the chromosomes

to be called chromosomes (coloured bodies). It was further observed that cells from different species have different number of chromosomes. It was also discovered that when a cell divides, the number of chromosomes in the newly formed cells remains the same. That is, the daughter cells have the same number of chromosomes as the mother cell. This observation gave rise to the idea that chromosomes may be the carriers of the units of heredity.

The first hint of the real nature of chromosomes came from the work of the Swiss biochemist Johann Friedrich Miescher. In 1869, he made the remarkable discovery that the chromosomes were, in fact, chemical molecules. For his research he used a material that most of us wouldn't even like to touch– pus from infected wounds! By treating white blood cells separated from pus-soaked bandages collected from a hospital with a weak alkaline

solution, Meischer was able to isolate a unique chemical substance containing both phosphorus and nitrogen. He first named the substance "nuclein" because it seemed to come from cell nuclei. By the early 1900s, Miescher's nuclein was found to be a mixture of proteins and acids. We now know the acids as nucleic acids. There are two of their kind – deoxyribonucleic acid (DNA), found mainly in the nucleus and ribonucleic acid (RNA), found mainly in the cytoplasm of the cell.

But even after the chemical nature of chromosomes was revealed, there was nothing indicating that they were indeed the carriers of heredity. In the first four decades of the 20th century, many scientists believed that the protein in the nucleus carried the genetic code, and DNA was merely a supporting "scaffold." The Russia-born American chemist Phoebus Levene was the first to moot the idea of proteins found in the cell nucleus to be the carriers of heredity. His argument was simple. Proteins were known to be long-chain molecules called polypeptides, made up of smaller units called amino acids. With 20 different amino acids to build from, there could be millions of ways of combining them to build proteins. Even if four amino acids were taken at a time, the number of different possible combinations comes to more than 160,000. DNA on the other hand was made up of only four different kinds nucleic acids. Levene believed that the DNA chain had repeating units of four nucleotides in different combinations. With such limited options available, he found it hard to see how DNA could be "intelligent" enough to carry hereditary information. In his opinion proteins were "a better choice as the nuclear hereditary material." After all, the 20-amino-acid "alphabet" of proteins could be configured into many more information-carrying structures than the four-letter alphabet of DNA. Of course, we now know, Levene was

wide of the mark.

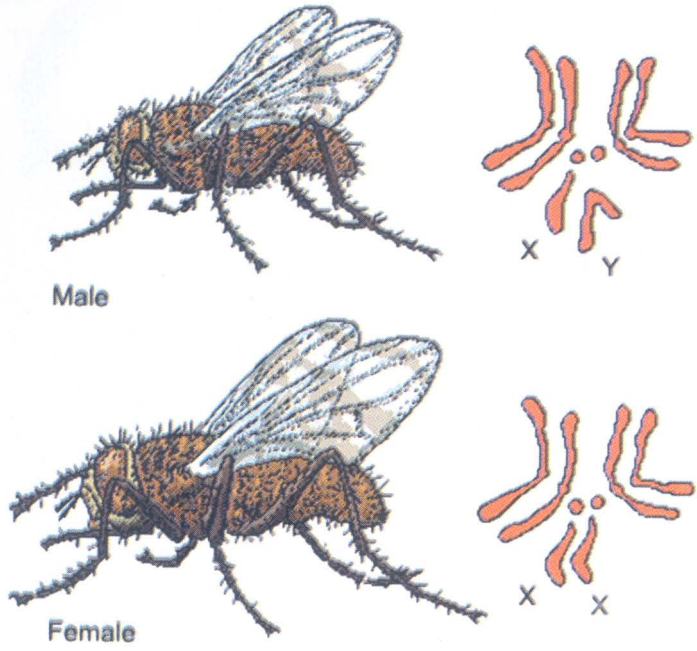
The idea that chromosomes were the real carriers of units of inheritance was first clearly pronounced by the American geneticist Walter Sutton in 1903, and by German cytologist Theodore Boveri in 1904. Working at Columbia University with grasshopper chromosomes, Sutton provided the first conclusive evidence that chromosomes carry the units of inheritance and occur in "definite, distinguishable, and different pairs of like (or homologous) chromosomes." He also suggested that the behaviour of the units of inheritance during division of the chromosomes when forming sex cells (meiosis) is the physical basis of the Mendelian law of heredity. Sutton developed this hypothesis in *The Chromosomes in Heredity* and concluded that chromosomes contain genes and that their behaviour during formation of sex cells (which contain half the number of chromosomes) by meiosis is random. For example, all male cells carry the sex chromosome pair X and Y, whereas in females it is X and X. So when a male sex cell (sperm) is formed from a male cell, it carries either an X or a Y sex chromosome, but which sperm will carry which cannot be predicted. But a female sex cell (ovum) carries only the X chromosome. Sutton's work formed the basis for the chromosomal theory of heredity. So it was clear that the sex of an offspring is a totally random event. Boveri's work with roundworms done at the University of Würzburg proved that a single chromosome is responsible for a particular hereditary trait.

But nobody at that time had any idea as to how the hereditary traits were really transmitted by chromosomes. It was in this background that an American zoologist Thomas Hunt Morgan began his researches in heredity. For his experiments, Morgan chose the common fruit fly, *Drosophila melanogaster*. The fruit fly was an ideal

choice because it could be easily kept alive and bred in labs, and since it could produce a new generation about every twelfth day, as many as 30 generations could be produced in a year. Further, males and females of the fruit fly could be distinguished easily, and it had only four chromosomes, which made their study simple.

Thomas Hunt Morgan, whose research with fruit flies established the link between chromosomes and heredity

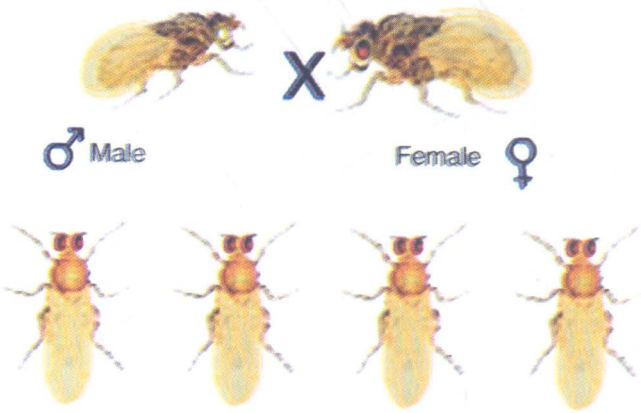
Morgan started his "Fly" lab at Columbia University in 1904 to study genetic variations in the fruit fly. However, unlike Mendel, who could easily find pea plants with distinctly identifiable traits, Morgan's team had to spend months searching for a fly with any unique trait that could be studied. At last, in January 1910 they found one – a male with white eyes. Fruit flies normally have red eyes. The white eye was obviously the result of some change in its genetic material caused by a process called mutation. When this mutant male was crossed with a normal female with red eyes, all the offspring of the first generation had red eyes. But when the first generation hybrid males were crossed with the first generation hybrid females, white-eyed flies were born in the Mendelian ratio of 1:3, although only the males had white eyes. When white-eyed males from the second generation were crossed with red-eyed females of the second generation, equal numbers of offspring with white and red colour were produced. Here, again, only males had white eyes. Since only males showed the white-eye trait, Morgan



Four chromosomes of the fruit fly

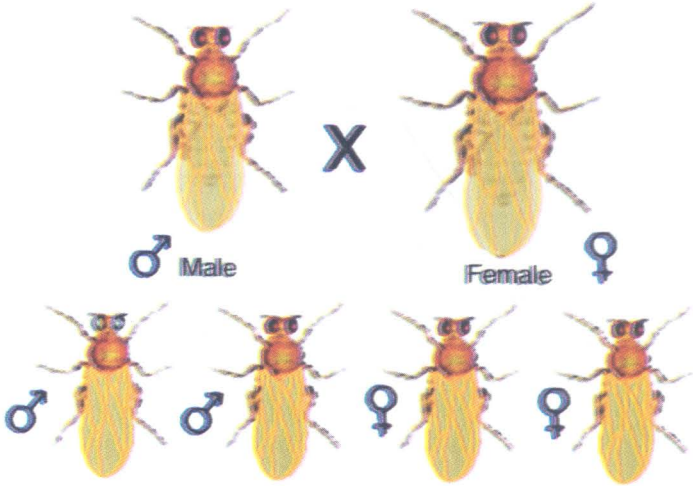
rightly concluded that this trait must be carried by the X chromosome, which is carried only by the male. This was the first recorded instance of a sex-linked trait.

Morgan's research with fruit flies brought out another significant feature of transmission of genetic traits. As Morgan and his group identified more and more inherited traits in fruit flies, they noticed that flies often showed particular combination of traits; that is, traits were not necessarily inherited independently, as Mendel had thought. It so happened that the seven characteristics of pea plants that Mendel had studies were governed by genes carried on separate chromosomes. In the case fruit fly, Morgan found that where two genes governing two different traits are located on the same chromosome, those traits are generally inherited together, just as a passenger



First generation (F1)

Crossing a white-eyed fly with a red-eyed one produced only red-eyed offspring



Second generation (F2)

When first generation red-eyed hybrids were crossed, white-eyed variety was born in the ratio 1:3

in the front seat of a car and one on the back seat travel together. But, just as a passenger can change cars, a piece of one chromosome occasionally switches to another, swapping places with a piece from it. Such crossing over usually occurs when the cell divides. Thus, Morgan's work provided further proof that genes were located on chromosomes.

In course of his research, Morgan found more than 80 different mutations in fruit flies. After years of painstaking studies, using both microscopic methods to study the chromosomes and Mendel's statistical method to analyse the transmitted traits, Morgan came out with four rules that governed the transmission of hereditary traits from parents to offspring. They are: (i) the combination rule; (ii) the rule of the limited number of the combination groups; (iii) the crossing-over rule; and (iv) the rule of the linear arrangement of the genes in the chromosomes. Morgan also prepared the so-called chromosome "map" of the fruit fly in which different hereditary factors could be located on the chromosome, like beads in a necklace. The chromosome maps provided definitive proof that Mendel's factors had a physical basis in chromosome structure. Later studies with lower plants and animals showed that, as a principle, Morgan's rules were applicable to all multi-cellular organisms. For his valuable contributions to the understanding of the hereditary process, Morgan was awarded the Nobel Prize for Physiology or Medicine in 1933.

2

It's in the Genes

When Morgan was working with his fruit flies, the concept of the gene was still nebulous. It was left to Hermann Joseph Muller, a member of Morgan's team, to finally identify the gene as a tiny cell organelle, which could be manipulated or even changed artificially. After leaving Morgan's group, Muller devoted his labours entirely to the difficult task of finding methods of changing the hereditary factors artificially. First, he created extremely elegant procedures by which mutation frequency could be measured exactly. It was known that the hereditary factors that reside in the chromosomes occasionally undergo spontaneous change, or mutation, which led to a change in certain traits passed on to the offspring. Mutations offered a simple way of identifying individual genes as any change in a gene could be easily linked to an observable change in a particular trait decided by the gene. But the frequency of spontaneous mutations was too low to arrive at any specific results through experiments. Muller hit upon the idea of artificially inducing mutation in the fruit fly using X-rays.

Hermann Joseph Muller, who pioneered artificial mutation using X-rays.

Experiments could then be arranged so that nearly 100 per cent of the offspring of the irradiated fruit flies showed mutations.

This idea that mutations could be artificially induced aroused a great sensation when it was published in 1927 and rapidly led to a great deal of work of different kinds in the most varied directions. It led to an amazingly rapid development in research in genetics. More importantly, since the effect of irradiation on the genetic material in living organisms was found to be universal, for the first time it led to the realization of the risks to human life from excessive exposure to X-rays and other ionising radiations. The extended knowledge of the mechanism of the mutation processes also triggered work in numerous fields outside theoretical genetics, such as plant improvement and research on evolution. For his pioneering studies on induced mutation, Muller was awarded the Nobel Prize for Physiology or Medicine in 1946.

The credit for identifying DNA as the key substance of heredity goes to three American biochemists Oswald Avery, Colin MacLeod and Maclyn McCarty. In the 1930s, they started their research with the pneumococcus bacteria that causes pneumonia. Two forms of the bacteria were known – one that caused the disease and another that was harmless. The infectious strain had a capsule-like coat made of sugars while the non-infective strain did not have any coat. Experiments carried out by a British microbiologist named Frederick Griffith had earlier shown that the harmless strain of the bacteria could become infectious when mixed with the infectious variety that had been killed by heat. This was quite surprising; how could dead bacteria pass on their infectivity to harmless bacteria? It was obvious that the dead bacteria were giving out some substance that was

"transforming" the harmless bacteria into infectious ones.

For a long time, there was a debate about whether bacteria had genes and what attributes they might have in common with higher life forms. To find that out, Avery's team at the Rockefeller Institute followed up Griffith's work with meticulously designed experiments, and discovered that the so-called "transforming principle" that transformed harmless bacteria into infectious ones was nothing but pure DNA. They used a very clever trick; they treated the killed infectious bacteria step by step, first with an enzyme that ate up the sugar coat, then with a protein-digesting enzyme and tested the remaining material. It could still transform. Thus the sugar coat and the proteins in the nucleus did not make up the transforming principle. But when they treated the killed bacteria with DNA-digesting enzyme, the transforming ability was destroyed. That is, without DNA the infectivity could not be passed on. So it was clear that bacteria did have DNA and that the DNA indeed carried the hereditary factors. The results of these classic experiments were published in 1944.

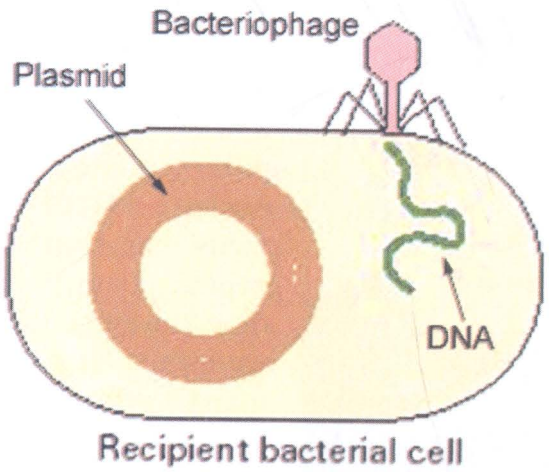
It was not only biologists who played the leading part in unravelling the DNA. Max Delbrück was an up-and-coming German theoretical physicist working in Berlin. In August 1932, while on a visit to Copenhagen, he heard a talk on "Light and Life" by the Danish physicist Niels Bohr. In his lecture, Bohr shared his thoughts on the phenomenon of life, taking into account

Max Delbrück

the latest breakthroughs in physics. Despite his complete ignorance of biology at the time, Delbrück was so impressed by the talk that he "instantly and firmly resolved to devote himself to biology". Back in Berlin, he met a Russian geneticist named Nikolai Timofeeff-Ressovsky, who "would speak for hours on end, explaining his science – genetics." He spoke about the mathematically precise laws of Mendel, about Thomas Morgan's remarkable work that proved that genes were arranged in chains in chromosomes. After listening to the discourses, Delbrück found that the most striking thing about genetics was its remarkable similarity to quantum mechanics, which had introduced into physics the concept of discreteness – the notion of "jumps" – and also had forced recognition of the role played by chance. The genes postulated by biologists were also discrete entities that also acted in a random manner.

The similarity between quantum mechanics and genes had a profound impact on Delbrück. He fell "captive to the great mystery locked up in the short word gene". "How do genes double, or reduplicate, when a cell divides?" he wondered. As luck would have it, around the same he was pondering over the nature of genes, Delbrück happened to learn of the existence of bacterial viruses, the so-called bacteriophages. These mysterious particles, which could hardly be described as living organisms, behaved just like large molecules when taken outside the cell – they could even form crystals. But in a short while after such a particle had penetrated a bacterium, the bacterial cell would burst open, and as many as 100 exact copies of the original virus would spill out. Delbrück could at once see the advantage of using bacteriophages to study the process of gene replication.

In 1937, Delbrück left Germany and went to the



Bacteriophage inserting its DNA into bacterial cell.

United States on a fellowship from the Rockefeller Foundation. The foundation's manager, Warren Weaver personally invited Delbrück to devote himself entirely to the study the problem of bacteriophage replication. Once in the US, Delbrück gathered a handful of young researchers to study the problem. Among them were Alfred Hershey and Italy-born Salvador Luria. The three scientists knew that bacteriophages had a simple structure – a coil of nucleic acid surrounded by a protein shell. When a bacteriophage infects a bacterial cell, it injects the nucleic acid into the cell by a simple but extremely efficient mechanism, while its protein shell remains outside. Delbrück, Hershey and Luria discovered that as a result of infection both the virus and the cell underwent drastic changes. The so-called cell-virus complex behaved as an essentially new system. The chemical activities of the cell were reprogrammed, finally leading to the formation of hundreds of new virus particles. In other words, the genetic make-up of the bacteria was entirely transformed by the infecting virus.

Alfred Hershey

The three scientists worked out quantitative methods to describe how phages multiplied and also demonstrated that the nucleic acid (DNA) from the virus was responsible for genetic changes in succeeding cell generations. They were thus able to demonstrate the role of nucleic acid as the carrier of genetic information of the virus. These findings were important in that it enabled others to develop vaccines against such viral diseases as mumps, polio and German measles. Delbruck, Hershey and Luria were awarded the Nobel Prize for Physiology or Medicine in 1969 for their discoveries concerning the replication mechanism and the genetic structure of viruses.

But the ubiquitous DNA still held a lot of secrets. That hereditary characters were transmitted from parents to offspring via special elements, the so-called genes, in the ovum and spermatozoon was known. It was also known that the organism that develops from the fertilized ovum receives certain of the parents' characters through these genes and that all the

Salvador Luria

genes combined determined the development of the organism. But the exact mechanism of how the genes actually transmitted the hereditary characters to the offspring still remained a mystery. It was still not known how the genes strung on the DNA molecule controlled protein synthesis that eventually manifested in the form of different traits.

The first step towards unravelling genetic control of protein synthesis was taken by American biologists George Beadle and Edward Tatum in the late 1930s. They conducted a series of cleverly designed experiments by using artificial mutation caused by X-rays. Working at Stanford University, Beadle and Tatum used the common pink bread mould, *Neurospora crassa*, to explore genetic changes caused by mutation. The method they used was quite elegant.

George Beadle

In the lab, *Neurospora* grows well on a simple culture medium that contains only a few simple sugars, inorganic salts and a vitamin. Obviously, the mould must be producing enzymes that convert these simple substances into the amino acids and other vitamins necessary for growth. The scientists reasoned that if they knocked off some of the genes by exposing the mould to X-rays, they should get strains that would not grow in the simple culture medium because once a gene was damaged the enzyme it coded for would not be produced. Thus, moulds exposed to radiation, they

Edward Tatum

reasoned, should lose the ability to produce certain essential nutrients, and this should slow, even stop the growth of the mould in a simple medium. If, however, the missing nutrients were added to the culture medium, the mutants should grow.

When they carried out their experiment, Beadle and Tatum found that this indeed was the case. By adding single nutrients to the culture medium and studying the growth for each radiated sample, they were able to conclude that a single gene determined the production of a specific enzyme that, in turn, allowed a single chemical reaction to proceed. Thus, the two scientists succeeded in demonstrating that the body substances (chemicals) are synthesized in the individual cell step by step, in long chains of chemical reactions, and that genes control these processes by individually regulating definite steps in the synthesis chain. They further showed that this regulation takes place through formation of special enzymes by the gene and that each gene produced one specific enzyme, not more. This came to be known as the "one gene–one enzyme" concept. This discovery was the key to our understanding of the manner in which the genes work, and has now become one of the foundations of modern genetics. Beadle and Tatum's technique has become one of the most important tools for the study of cell metabolism and has yielded results, which are of great

significance to various problems in the fields of medicine and general biology.

Joshua Lederberg

Tatum achieved another breakthrough in 1946 when, working with fellow geneticist Joshua Lederberg, he was able to transfer genes from one organism to another artificially. Working with bacteria, which reproduce asexually, they found that the mixing of two different strains of a bacterium resulted in genetic recombination between them and thus to the creation of a new, crossbred strain of the bacterium. In bacteria, genes are exchanged through a "mating channel" that links two bacteria during conjugation; in higher life forms genes are passed on to subsequent generations by parents following sexual union.

Lederberg's team also discovered that bacteriophages were capable of carrying a bacterial gene from one bacterium to another – a phenomenon they termed transduction. They also showed that if bits of genetic material from another organism are introduced into the bacterial body, like spliced tapes, they become part of the genetic material of the bacterial cell and thus change its constitution. This was the first example of experimental manipulation of an organism's genetic material by introducing new genes into it. For their outstanding contributions to our understanding of the very basis of heredity and the manner in which the genes function, Beadle, Tatum and Lederberg were jointly awarded the Nobel Prize for Physiology or Medicine in 1958.

As the genetic mystery was unfolding, it became evident that in all living organisms – be it viruses, bacteria, plants, or animals – proteins and nucleic acids are always present as the life-supporting elements. It was also known that both proteins and nucleic acids are very large molecules, built up from smaller units linked together in chains – just like strings of pearls – which often form helices. As we know, all proteins are polypeptides made up of combinations of only some twenty amino acids, while nucleic acids are built up of molecules called nucleotides – made up of nitrogenous bases, sugar and phosphoric acid. There are no more than eight of these most important nucleotides found in living organisms.

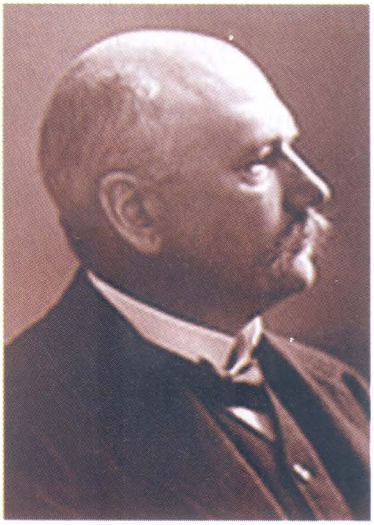
All nucleotides contain a sugar and a phosphate group, but only one of five different kinds of nitrogenous base. The sugar can be of two kinds – one of which, called ribose, contains one more oxygen than the other, called deoxyribose. Interestingly, it is this seemingly insignificant difference in a single atom that produces a remarkably great effect, giving rise to two distinct kinds of nucleic acids – ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) – which have widely different functions in a cell.

Each molecule of DNA is enormously long. For example, the DNA in every cell of the human body would measure almost two metres in length and the total DNA in any one of us would stretch more than 960 million kilometers; that is, more than six times the distance between the Sun and the Earth!

Despite its enormous length, however, the DNA molecule is built on a very simple plan. It features a backbone, in which the same pattern of atoms occurs over and over again, in some cases as often as several million

times. Two alternate groups – the sugar deoxyribose and a phosphate – are joined together to give the sequence: phosphate-sugar-phosphate-sugar-phosphate-sugar, and so on. Thus the backbone is very regular and always has the same pattern. As it turned out, the difference is only in the nitrogenous bases attached to each sugar, and it is the order of these bases that gives DNA its unique character as a molecule of heredity.

Although their real nature was revealed much later, work on nucleic acids had been going on for quite some time. As early as 1879, the German scientist Albrecht Kossel began studying the newly isolated substances (which were dubbed "nucleins" by the Swiss biochemist Johann Friedrich Miescher) and by careful hydrolysis showed that they consist of a protein portion and a non-protein portion, namely nitrogen-containing nucleic acids. He and his students then used various techniques to chemically analyse the nucleic acids, thus discovering their component compounds: the nitrogenous bases adenine, cytosine, guanine, thymine, and uracil. For his elucidation of the chemistry of the nitrogenous bases of the nucleic acids, Kossel was awarded the Nobel Prize for Physiology or Medicine in 1910.



Albrecht Kossel

By the late 1940s, the English chemist, Alexander Robertus Todd had been able to build up a variety of nucleotides out of simpler fragments and describe in detail the chemical properties of the nucleic acids. But

no one at that time ever tried to synthesize nucleic acids because they were considered too complex, being made up from 100 to 10,000 nucleotide units in each molecule. By this time, however, it was known that DNA was a polymer with a backbone of sugar and phosphate linked to nucleic acids, but the molecular structure of DNA still appeared a distant dream.

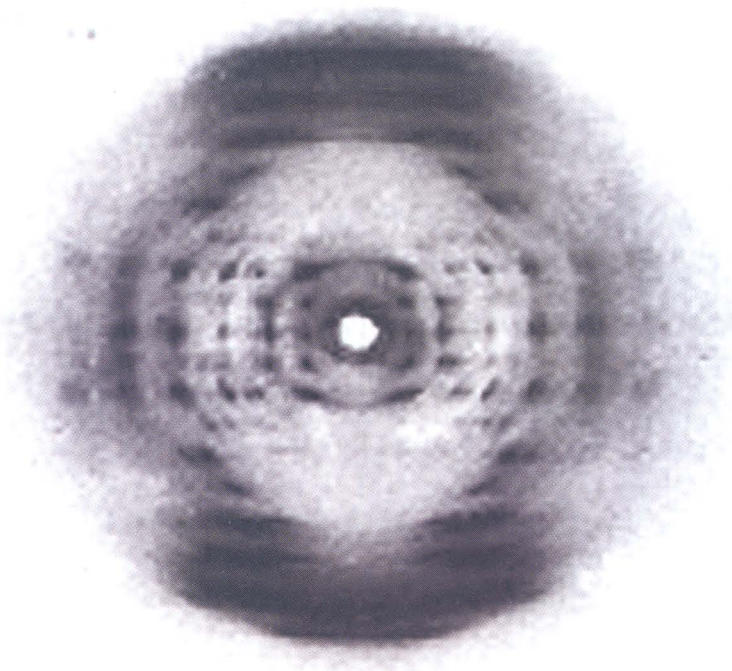
Another vital clue about DNA came from the work of Austria-born American biochemist Erwin Chargaff, who isolated DNA from different organisms and tried to find out the amount of the nitrogenous bases in each. To his great surprise, he found that the amount of thymine in DNA was always the same as the amount of adenine. Not only that, what was more surprising was that the amount of guanine was always the same as the amount of cytosine. Chargaff's findings immediately ruled out the possibility of DNA being made up of groups of four nucleotides, as proposed by Phoebus Levene, because in that case the amounts of all the four nucleotides should have been equal. Chargaff's 1 : 1 ratios of A : T and C : G provided the key evidence which finally led to the solving of the DNA puzzle.

3
The Double Helix

By the early 1950s, the chemical nature of DNA was well established. It was known that DNA is composed of building blocks called nucleotides consisting of a deoxyribose sugar, a phosphate group and one of four nitrogenous bases – adenine (A), thymine (T), guanine (G), and cytosine (C). It was also known that DNA contained equal amounts of A and T on the one hand and equal amounts of C and G on the other. Further, it was known that in the DNA molecule the nucleotides were linked in a series – from one phosphate to the next sugar, to the next phosphate, to the next sugar, and so on.

But none of the early researches on DNA could explain the mechanism by which the hereditary traits were actually transmitted from one generation to another; that is, how the genetic information was passed on by DNA. To do that it was essential to find out how the nucleotides were arranged in the three-dimensional DNA molecule. It was obvious that whatever the structure of DNA, it ought to be able to explain the basic function of the genetic material – that of self-copying, or replication – which was essential for transmission of the hereditary characteristics from one generation to the other. Analysis of X-ray diffraction patterns of DNA provided the vital clue.

Scientists use various techniques to decipher the molecular structure of substances. One of the techniques they use is X-ray diffraction, in which a narrow beam of



X-ray diffraction pattern of DNA

X-ray is passed through a sample to produce a "diffraction pattern" in the form of symmetrically arranged spots of different intensities on a photographic film. The resulting diffraction pattern is a unique "signature" of the molecule. By analysing the diffraction pattern the geometry of the X-ray scattering atoms, or in other words, the molecular structure of the sample material can be worked out.

As early as 1947, the British biochemist William Astbury had published one blurry X-ray diffraction photograph of DNA; it revealed hardly anything. Soon after, Sven Furburg, a Norwegian student in London, took a few more diffraction pictures and deduced that

DNA's four bases lay at right angles to the sugar-phosphate chain. He reported his results in his doctoral dissertation and returned home; nothing more was heard of his work. Then, in 1951, working at the California Institute of Technology, the famous American chemist Linus Pauling used the X-ray diffraction technique to work out the structure of proteins, which were found to be shaped like a coiled spring – the so-called alpha helix.

Linus Pauling

In Cambridge, England, the New Zealand-born British biophysicist, Maurice Wilkins had been studying X-ray diffraction patterns of DNA made by fellow scientist Rosalind Franklin, which were much better than the earlier results, and had collected a vast amount of data on the structure of the molecule. On the basis of the X-ray patterns he had postulated that the DNA molecule had a helical shape (because their diffraction patterns look like the diffraction patterns of proteins, which, by then, were known to be helical shaped). He also succeeded in measuring the approximate diameter of the helix.

Maurice Wilkins

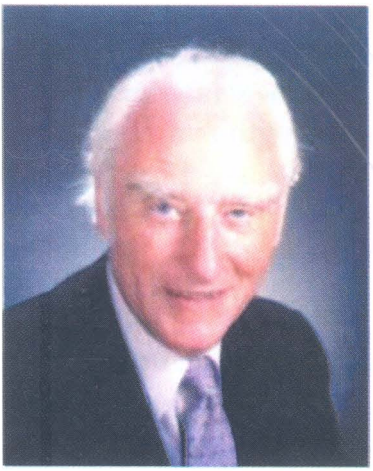
But Wilkins' data did

not give much information about how the DNA chain was arranged within the helix or how the nitrogen bases were linked. That revelation came from the work of two young Cambridge scientists, James D. Watson and Francis Crick. Watson had met Wilkins at Naples, Italy in 1951, and had become acquainted with the X-ray diffraction patterns of DNA molecules. On his return to Cambridge, he met Crick, with whom he began working to solve the puzzle of the structure of DNA. After studying the data accumulated by Wilkins on X-ray diffraction, Watson and Crick took little time to deduce that DNA should be a double helix with the phosphate groups on the outside and the bases on the inside. From the measurements made by Franklin and Wilkins they were also able to deduce the basic dimensions of the helix.

Rosalind Franklin

Watson and Crick were also influenced by the work of Avery, who had conclusively shown that DNA was the substance responsible for heredity. They knew DNA had to be "intelligent". It occurred to them that if the order of the nucleotides changed, information could be coded into the DNA sequence, much like the Morse code for alphabets used in telegraphy. With four different bases, the combination of sequences possible could be enormous. They believed, "DNA, not protein, was the Rosetta Stone for unraveling the true secret of life."

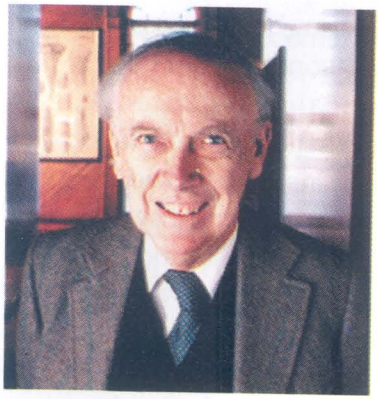
But the structure of DNA was not solved yet.



Francis Crick

Although Watson and Crick had some idea of the basic helical structure of the molecule, big questions still remained unsolved. It was still not clear how the two helices were joined together nor was it known how the bases were arranged. And time was running out. Watson and Crick were aware that after solving the alpha-helix structure in proteins, Pauling at Caltech would be trying to solve the DNA structure, which he almost did. Even as the Cambridge duo was trying to work out the DNA model, the news came that Pauling had already submitted a paper on the structure for DNA. It came as a shock, but after seeing Pauling's paper, Watson and Crick realised that the proposed model wouldn't work, simply because it was a triple-helix that could never account for replication, so essential for it to function as a carrier of heredity.

Pauling had put the phosphate at the core of each helix with the bases on the outside and this was a major flaw in his model. He had overlooked the fact that the negative charges of the oxygen atom in each phosphate group in the middle of the proposed model would repel one another, making it impossible for the molecule to hold together. Watson and Crick were, however, aware that Pauling would notice his mistake and redouble his efforts to set it right, and that they had no time to be complacent. So, not long after Pauling's paper came out, Watson began to play with paper cutouts of the nitrogen bases. He knew that nucleotides could pair and form



James Watson

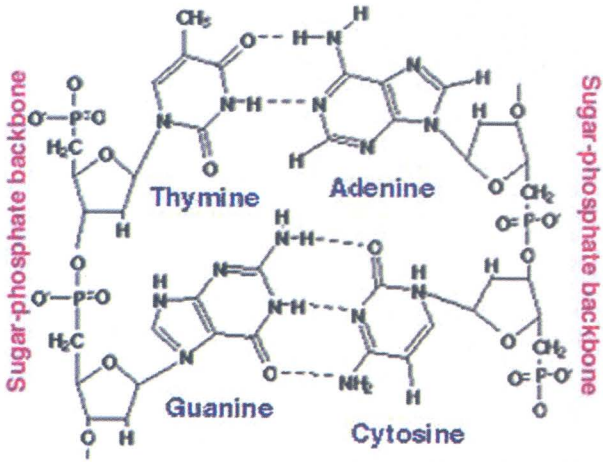
weak bonds called hydrogen bonds. Now the question was which bases paired with which ones.

Of course, it was known that two of the nitrogenous bases were small single-ring molecules called pyrimidines and two were larger, two-ring molecules called purines. If the diameter of the DNA double helix were uniform throughout, it was obvious that one of the smaller base molecules must be bonding with one of the larger ones. But the question was, which was bonding with which? Here, Chargaff's work came in handy. Chargaff had found that in DNA, the amount of adenine always equalled that of thymine and the amount of cytosine always equalled that of guanine. Could it be that adenine always paired with thymine and cytosine always with guanine? Using the paper cutouts and "in a moment of part insight and part luck", Watson was soon able to confirm that it indeed was so.

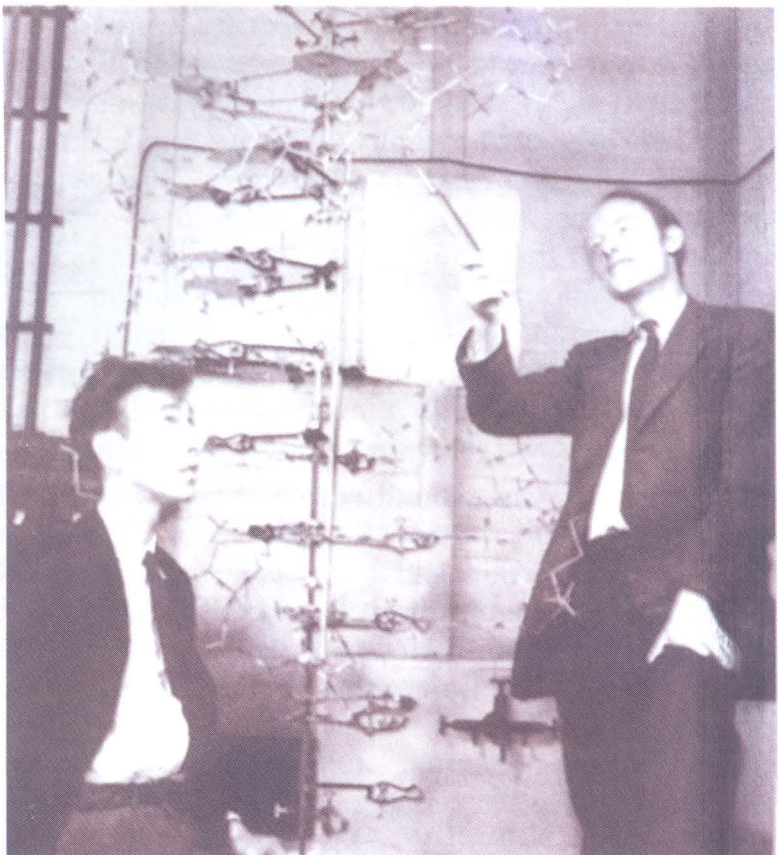
Comparing the width of different hydrogen bonded pairs of bases using his paper cutouts, he found that some pairs had vastly different widths. If these pairs really occurred in the DNA helix then the helix would be uneven and would bulge in and out. But X-ray diffraction pattern of DNA shows it to be a helix of uniform width. If, however, the pairings were between A-T and C-G, the two base pairs had exactly the same width and so would make the helix of uniform width. Further, this scheme of base pairing also agreed with Chargaff's ratios, and allowed the bases to compactly stack on top of one another. When Watson showed his scheme to Crick, he

immediately agreed. They were convinced that base pairing was the key to the DNA structure, which indeed it was. Crick further pointed out that "because of certain bond angles and the proximity of the base pairs, the two helices had to run in opposite directions". That is, the sequence of bases in one strand travels in an opposite direction from the sequence of bases in the other. In scientific parlance, the helices are said to be "antiparallel" to one another. On Saturday, February 28, 1953, Watson and Crick walked into the "Eagle," a dingy pub in Cambridge, England, and announced to the lunchtime crowd that they had discovered the secret of life.

The next step was to build a three-dimensional model using metal scraps and wire. On March 7, 1953, Watson and Crick first pieced together a six-foot (180-cm) model of DNA, incorporating the A-T and C-G base pairing scheme and the Crick's idea of antiparallel strands. It was an elegant model in which everything fitted into



Base pairing between guanine-cytosine and adenine-thymine



Watson and Crick with their model of DNA

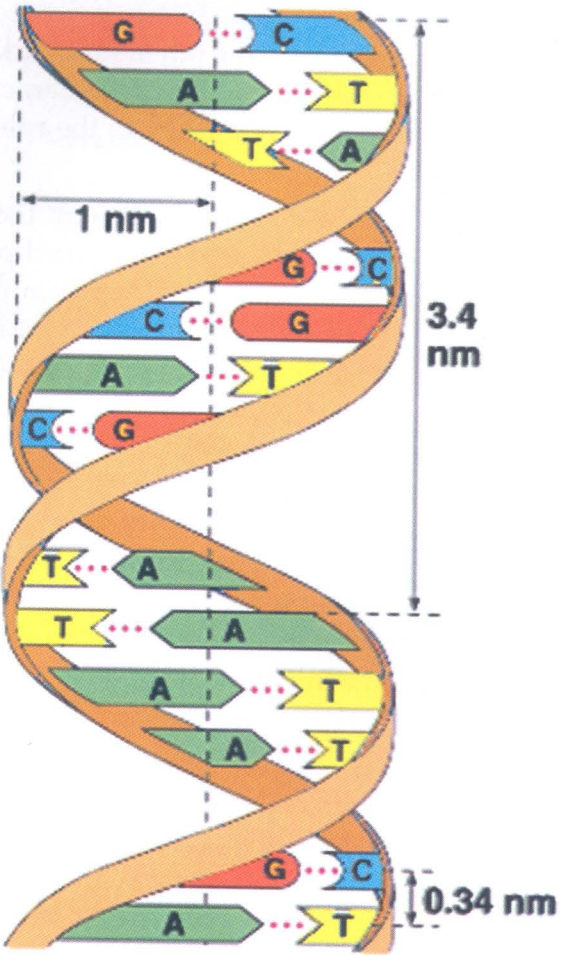
place beautifully. When Wilkins and Franklin were shown the model, they too agreed that it was the right DNA structure. So, at last, a plausible structure of DNA was in place. It resembled a twisted ladder, where the sugar and phosphate form the rails, and the base pairs are the rungs. The rails run in opposite orientation to each other.

The unique thing about the Watson and Crick's DNA model is the complementary nature of the opposite rungs. For example, wherever there is an A on one strand, there is a T in the same position on the other strand. Similarly,

wherever there is a C on one strand, there is a G in the same position on the other strand. So, if the sequence of bases in one strand is known the other strand can be built up automatically. This complementary character of the base pairs, as we shall see, is the key to the role of DNA as a carrier of heredity.

Watson and Crick were so excited with their model that they lost no time to write up a 900-word paper and send it to the British journal *Nature*, which published it in its issue dated April 25, 1953. "The novel feature of the structure," Watson and Crick wrote in their paper, "is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from the other chain, so that the two lie side by side with identical z-coordinates. One of the pair must be a purine and the either a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6".

"If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations), it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine). In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined."



The DNA double helix

Watson and Crick further added, "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material." This was perhaps the biggest understatement in the history of science!

Looking back, we find that even before Watson and Crick's work, the DNA molecule was fairly well understood. Other researchers had shown that the DNA molecule is a high polymer chain composed of a few types of building blocks, which occur in large numbers. It was also known that in DNA these building blocks are a sugar, a phosphate, and a nitrogen-containing chemical base. The same sugar and the same phosphate are repeated throughout the giant molecule, but with minor exceptions there are four nitrogenous bases. The importance of the work of Watson and Crick lay in their pinpointing the way the bases were linked in the DNA molecule. They were the first to propose that the DNA molecule was a double helix, with the two helices joined by pairs of nitrogenous bases, with adenine always pairing with thymine and cytosine always pairing with guanine. It is this specific paring of these bases in the DNA double helix that makes it unique as an agent of transmission of heredity.

The proposed structure was so profound that it could immediately explain the key property of DNA – that of transmission of heredity by replication. During replication, Watson and Crick contended, the double helix unwound, opening up the nitrogenous base pair links like a zipper. Once open, the bases again paired off – adenine to thymine and cytosine to guanine – building up two complementary chains, which finally ended up creating two identical DNA double helices. Subsequent experiments have confirmed the accuracy of this model of the DNA molecule. For their discovery of the three-dimensional molecular structure of the DNA, Watson, Crick and Wilkins were awarded the Nobel Prize for Physiology or Medicine in 1962.

When Watson and Crick built their 3-dimensional model of DNA, there was no means available for actually

... April 25, 1953 NATURE 737

...equipment, and to Dr. G. L. R. Donovan and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

...

...

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

...

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagrams). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Pauling's model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Pauling's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If A is assumed to occur only in the form in which it has recently been found to occur, namely with the keto rather than the enol configuration, it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[2] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

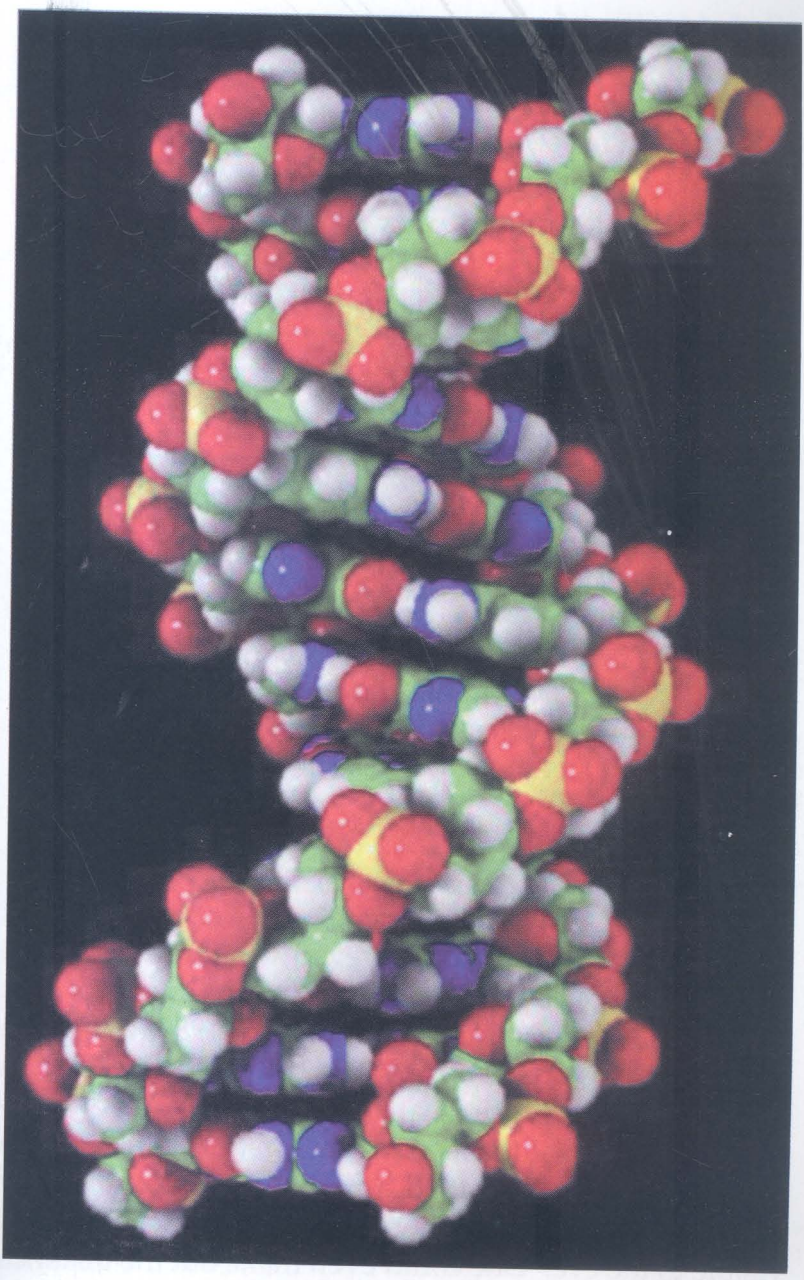
The previously published X-ray data[3,4] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

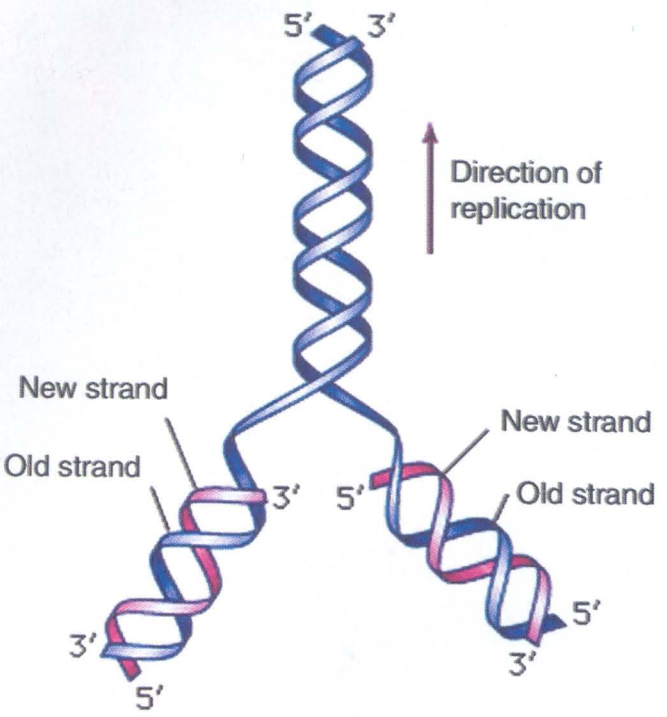
Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

The first paper on DNA structure by Watson and Crick, which was published in the journal *Nature* on April 25, 1953



3-dimensional model of DNA



Replication of DNA

observing the real thing. The molecule was so small that it would be invisible under even the most powerful microscopes. The first opportunity to have a direct look at the DNA molecule came in January 1989. Using the newly discovered "scanning tunnelling electron microscope" scientists at the Lawrence Livermore Laboratory in California took the first ever look at the DNA molecule, magnified a million times by the new technological marvel. There it was – the double helix, visible to the eye! One could even make out the distance between successive coils of the DNA helix. It was a remarkable achievement.

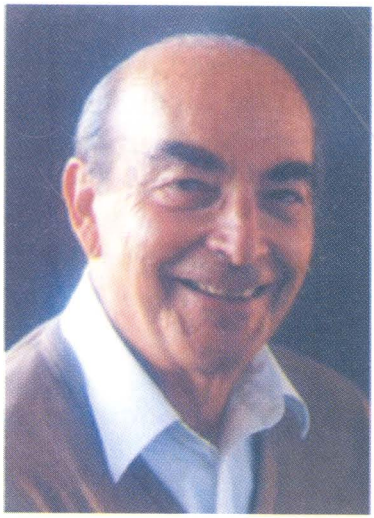
Watson and Crick's 1953 model of the double-helix structure of DNA not only provided an icon for a new generation of life scientists, but its latent potential also helped generate large funding for research in genetics that led to far reaching advances in biotechnology. As fifty years of explosive study has shown, understanding DNA changed everything. Within a decade, it explained exactly how genes stored genetic information in a four-letter chemical alphabet, how they replicate, and how they drive not only the development and functioning of the cell, but heredity and evolution as well. Watson later matched his scientific intuition with an elegantly simple biography, *The Double Helix*, which not only tracked the duo's adventure in research, but also isolated a turning point in history of biology.

In the half century since Watson and Crick worked out the structure of DNA, research into the double helix has brought about a revolution in our understanding of development, function, and disease in living organisms, especially humans. It has also become a commonplace truth that DNA (or RNA) is the unique *sine qua non* of any living organism. DNA is now being used as an identifier, a collection of base-pair sequences that provide an individual's unique identity.

4
Copying from the Old

The discovery of the double-helix structure of the DNA was only the first step in our understanding of the mechanism of transmission of genetic traits. The nuts and bolts of the actual process of transmission of genetic traits from cell to cell and the role of other players in the game still remained to be worked out. In 1954, soon after Watson and Crick's discovery of the DNA structure, the Russia-born American physicist and writer George Gamow suggested that the sequence of bases in DNA might act as what we now call a genetic code that regulated cell processes, making an organism what it is. But exactly what the code was and how it worked remained a mystery. In fact, two separate mechanisms had to be accounted for: (i) How the DNA molecule replicated itself; and (ii) How the DNA code was transformed into proteins.

When they proposed the double helix structure, Watson and Crick had suggested that during replication the two strands of DNA would unwind and separate and, because they are complementary to one another, each strand would act as a template for synthesising a new complementary strand. In fact, it was the possibility of complementary bonding of the bases in the two DNA strands, that led Watson and Crick to propose a double-helix structure rather than a triple helix, as proposed by Pauling. It was immediately clear that the synthesis of two complementary strands would lead to the formation of two identical double helices, one strand in each of which would be an "old" strand from the parent



Arthur Kornberg

molecule, and the other a newly synthesised strand.

The first confirmation of this hypothesis came from the work of the American geneticist Franklin W. Stahl and fellow molecular biologist Mathew S. Meselson, who showed that in a cell, new DNA is made by copying from the old, as proposed by the Cambridge scientists. By a series of ingenious experiments, using bases containing radio-labelled nitrogen isotopes and following their uptake by successive generations of bacteria in cultures, Meselson and Stahl showed that one strand of each DNA molecule is passed along unchanged to each of two daughter cells, just as proposed by Watson and Crick. Meselson and Stahl published their results in 1958.

Further proof of the complementary nature of the two strands of DNA came from the work of the American biochemist, Arthur Kornberg, than director of the microbiology department at Washington University, Kornberg concentrated on the study of the way living organisms make nucleotides. He then went on to find out how the nucleotides are strung together to form DNA molecules. In 1956, during course of this research he isolated and purified an enzyme (now known as DNA polymerase) from the common *Escherischia coli* bacterium. This enzyme, in combination with certain nucleotide building blocks, could produce precise replicas of short stretches of DNA in a test tube. Kornberg further showed how, during replication, chains of DNA are built up in

the cell, with complementary nucleotides adding on to the template provided by a half-DNA molecule. He also found that replication takes place only when intact DNA template and all the four nucleotides are present. If the DNA template is fragmented or if only one of the four nucleotides were missing, no DNA strand was synthesised.

Severo Ochoa

Meanwhile, in 1955, the Spain-born biochemist Severo Ochoa had discovered an enzyme in bacteria that could catalyse the formation of RNA from nucleotides. He named it "polynucleotide phophorylase". Surprisingly, although under test-tube conditions this enzyme was found to help synthesise RNA, in its natural environment in the cell, its function is to degrade RNA. Nevertheless, the enzyme has been singularly valuable in enabling scientists to understand and re-create the process by which the hereditary information contained in genes is translated, through RNA intermediaries, into enzymes that determine the functions and character of each cell. For their valuable contribution in unravelling the mechanism for the biosynthesis of DNA and RNA, Kornberg and Ochoa were awarded the Nobel Prize for Physiology or Medicine in 1959.

As we all know, DNA, with its genetic codes, is found mostly in the cell nucleus, but protein synthesis takes place outside the nucleus, in the cytoplasm. So the next

question that needed to be answered was: How does DNA direct the synthesis of protein from inside the nucleus? DNA is like a blueprint; it has to be "read" and the instructions passed on to some agent that would put the amino acids together to build up the specific protein. This is like building a house, where the civil engineer reads the blueprint and directs the masons, who, in turn build the house. Crick mooted the idea of a carrier molecule that carried the information from DNA in the cell nucleus to the cytoplasm for protein synthesis. He suggested that the other kind of nucleic acid, namely RNA, which is commonly found in the cell's cytoplasm, could be an ideal candidate for this role. In brief, Crick proposed that the route to the synthesis of protein on the basis of the genetic code carried by DNA was "from DNA to RNA to protein." This scheme came to be known as the Central Dogma. According to this scheme, the genetic code is first copied to RNA by the process known as transcription; the information from RNA is then used to make protein by the ribosome in the cytoplasm through the process of translation. As we shall see later, Crick was on the right track.

Like DNA, RNA also has a sugar-phosphate backbone, but the sugar in RNA is ribose rather than deoxyribose. Another difference is that RNA is not double-stranded but a single-stranded molecule in which the nitrogenous base thymine is replaced by the uracil (U). But that does not create any problem. Since uracil is also a pyrimidine base and can join with adenine just like thymine, its presence does not prevent RNA forming a complementary strand with half-DNA during transcription. So, the carrier RNA could easily copy the hereditary information from DNA by making a complementary strand.

George E. Palade made another important discovery. By careful study of electron micrographs of cells he discovered tiny particles in the cytoplasm that he demonstrated were the sites of protein synthesis. These particles, about 2 millionths of a centimetre in diameter, were rich in RNA and were therefore called ribosomes.

The next question that needed an answer was: How do the amino acids that ultimately make up the protein chain interact with the carrier RNA? In other words, how does the carrier RNA identify the amino acids to be brought together to form the polypeptide chain? To account for this, Crick proposed the existence of an "adaptor" molecule for each of the 20 amino acids that read the genetic code and selected the amino acid to be added to the growing polypeptide chain. Crick's adaptor molecule later turned out to be very low-molecular-weight RNAs that were soluble in water. It was at first named "soluble" RNA, but was later renamed "transfer" RNA, or tRNA, to better reflect its role in ferrying amino acids to ribosomes for protein synthesis.

Francois Jacob

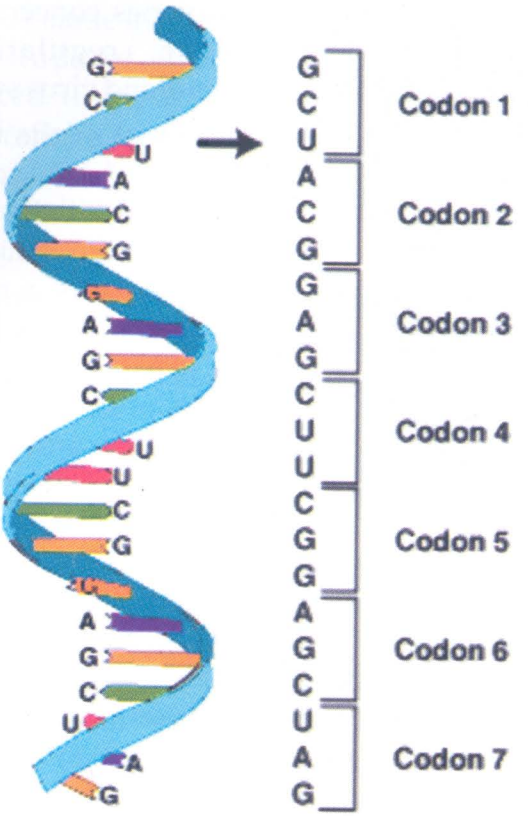
It was, however, clear that whatever way they did it, the genes could not be producing proteins all the time; there had to be some mechanism for regulating gene expression. First hints of how the genetic code in DNA regulates cell growth came from the work of three French scientists – Francois Jacob, Jacques Monod and André Lwoff, who made important

Jacques Monod

discoveries concerning the genetic regulation of enzyme and virus synthesis.

Most of the work of Jacob and Monod was carried out at the Pasteur Institute, Paris, which Jacob had joined as a research assistant in 1950. In 1958, they began to collaborate in studies of the regulation of enzyme synthesis in bacteria. They were the first to suggest the existence of a messenger RNA, or mRNA, a single-stranded molecule whose base sequence is complementary to that of DNA in the cell. They postulated that during translation of the genetic code, the messenger carries the "information" encoded in the base sequence in DNA to ribosomes, which are the sites of protein synthesis.

The proposed mechanism was quite simple. When the DNA molecule unwound, as suggested by Watson and Crick, one of its strands replicated its structure, not as sequence of nucleotides that form a DNA molecule, but as a sequence of nucleotides that form an RNA molecule. Thus the adenine of DNA strand does not pair with a thymine, but with a uracil. (This process is different from what happens in DNA replication, when two complementary strands of DNA are built up.) The resulting mRNA molecule, carrying the genetic code imprinted on its nucleotide pattern, then leaves the nucleus and passes into the cytoplasm where protein synthesis takes place. In the cytoplasm, the base sequence of the mRNA is translated into the amino acid sequence of an enzyme, which is basically a protein. Thus mRNA



Codons are sequences of three nitrogenous bases in mRNA that specifies a single amino acid

can be considered the "working copy" of the genetic information contained in the DNA "master copy" stored in the chromosome.

The next step was to find out how the genetic code instructed the tRNAs and amino acids to make protein. It was at first conjectured that ribosomal RNA, or rRNA, was the "template" on which proteins were built, but it was soon realised that rRNA did not have the right properties.

Today we know that there are three main types of

RNA: messenger RNA (mRNA), transfer RNA (tRNA), and ribosomal RNA (rRNA). During protein synthesis, mRNA carries the codes from the DNA in the nucleus to the sites of protein synthesis in the cytoplasm (the ribosomes). Ribosomes are composed of rRNA and proteins; they can "read" the code carried by the mRNA. A sequence of three nitrogenous bases in mRNA (called a codon) specifies incorporation of an amino acid; tRNA brings the amino acids to the ribosomes, where they are linked into polypeptide chains (proteins).

Jacob and Monod proposed their famous operon theory in the early 1960s. In their classic paper they described the regulatory mechanism of the *lac* operon of the bacterium *Escherichia coli.* When lactose is available, they found that the bacterium turned on an entire suite of genes to produce enzymes to metabolise the sugar. In absence of lactose, the same genes were turned off. When they tracked the events that the addition of lactose initiated, they found that lactose removed an inhibitor from the DNA that suppressed the action of the particular genes. By switching off some genes when not needed, the molecule conserved energy, much the same way we do by switching off lights and fans in an empty room. In advancing the concept of operons, the two scientists postulated the existence of a class of genes that regulate the function of other genes by controlling the synthesis of mRNA.

According to the operon theory, near each bacterial gene is a segment of DNA known as the promoter. This is the site at which RNA polymerase, the enzyme responsible for the production of mRNA, sticks to the DNA and starts transcription. Between the promoter and the gene there is often a further segment of DNA called the operator, where another protein—the repressor — can stick. When the repressor is attached to the operator,

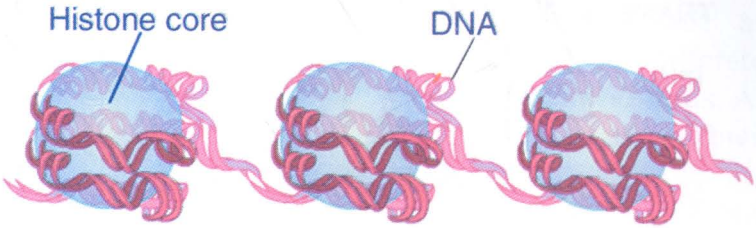
it stops the RNA polymerase from moving along the DNA chain and producing mRNA; consequently, the gene is made inactive. When the gene works, a single unit of mRNA is transcribed and is subsequently translated into separate proteins. A repressor protein was isolated in 1967 and found to be a rather small molecule. Thus the French scientists were able to demonstrate how the structural information of the genes was used chemically to synthesise proteins. Their discovery of a previously unknown class, called the operator genes, which control the structural genes, marked a major breakthrough.

Lwoff worked with viruses known as bacteriophages that infect only bacteria and found that after infection the genetic material of the virus is passed on to succeeding generations of the bacteria in a non-infective form. But, in some cases, the genetic material turns infective and causes disintegration of the bacterial cell, spilling out hoards of infective viruses in the process.

André Lwoff

Together, the work of the three – Jacob, a cellular geneticist, Monod, a biochemist, and Lwoff, a microbiologist – opened up a field of research which in the truest sense of the word can be described as molecular biology. For their work in unravelling the mechanism of gene action, Jacob, Monod and Lwoff were awarded the Nobel Prize for Physiology or Medicine in 1965.

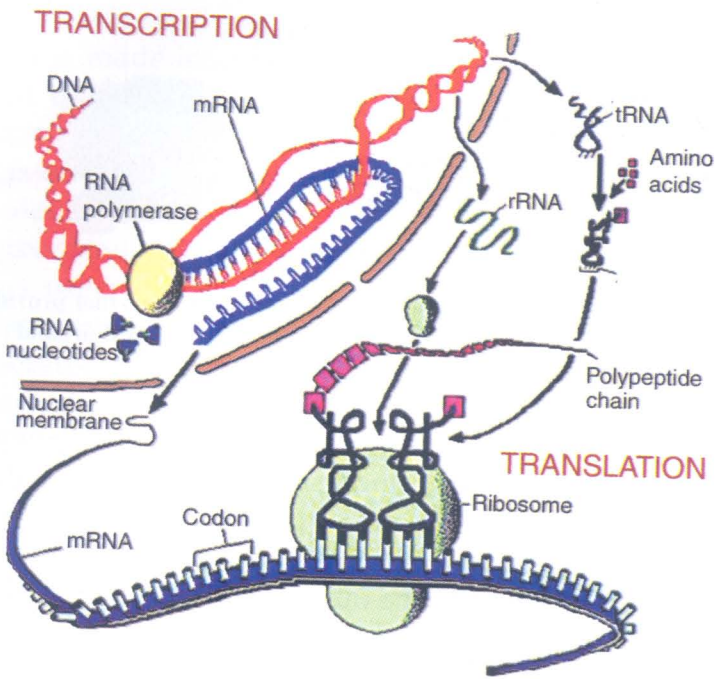
By the early 1970s, the overall mechanism of



Histones are structural proteins found in the cell nucleus that provide a scaffold upon which the DNA double helix is wrapped around

transmission of genetic traits by DNA of chromosomes became reasonably well understood. It was known that each chromosome is essentially a package of one very long, continuous strand of DNA. Up to two metres of DNA is packaged to fit the nucleus of one cell. In higher organisms, structural proteins called histones provide a scaffold upon which the DNA double helix is wrapped around. Five different kinds of histones have been identified. The DNA-histone complexes are then double-coiled (like the coiled coil filament used in incandescent tungsten lamps) in the shape of chromosomes. Thus histones play an important role in "packaging" DNA in chromosomes.

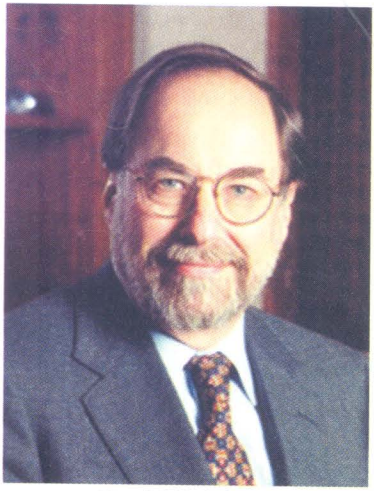
As is visible under a microscope, the cytoplasm of any cell is full of floating ribosomes, RNA polymerases, tRNA and mRNA molecules and enzymes, all carrying out their tasks independently of each other. The process of transfer of genetic information from DNA to protein takes place in two stages. When a gene is "on", the DNA unwraps from the histone scaffold and unwinds like a zipper. In the first stage, called transcription, RNA polymerase enzymes attach to the DNA strand at the starting points of different genes and copy the gene from



The steps in protein synthesis

DNA into an mRNA molecule. The mRNA molecule then floats over across the nuclear membrane to a ribosome, which reads the code and stitches together the string of amino acids into protein molecules it encodes, by the process of translation.

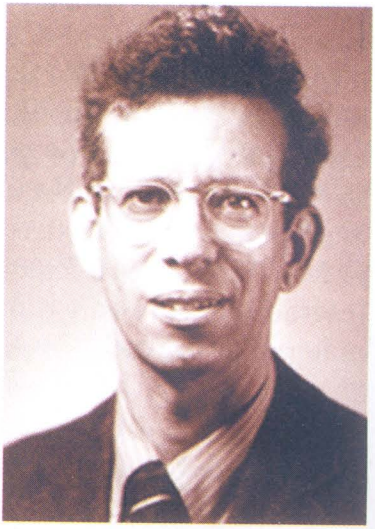
The site of protein synthesis, as we have already seen, is the ribosome – an extremely complex structure of proteins and rRNA bonded together into a large molecular machine. A ribosome is helped by a chemical called adenosine triphosphate (ATP), which powers it as it "walks" along the messenger RNA and stitches the amino acids together. It is also helped by tRNAs, a



David Baltimore

collection of 20 special molecules that act as carriers for the 20 different individual amino acids. As the ribosome moves down to the next codon on the mRNA, the correct tRNA molecule, carrying the correct amino acid, moves into place. The ribosome breaks the amino acid off the tRNA and stitches it to the growing chain of the polypeptide. The ribosome then ejects the "empty" tRNA molecule so that it can go and get another amino acid of the correct type. The same process works for all the codons on the mRNA.

Although applicable in most cases, the central dogma of molecular biology did not appear to hold in case of certain viruses. Two American virologists, David Baltimore and Howard Temin, and the Italian virologist Renato Dulbecco, while studying the transmission of genetic information from virus to cell stumbled upon a surprising fact. According to the central dogma, genetic information could be passed on only from DNA to RNA, and not *vice-versa*. But the work of Baltimore, Temin and Dulbecco showed that



Howard Temin

Renato Dulbecco

information could indeed be passed on from RNA to DNA in certain types of tumour-causing viruses. They also discovered that when a virus infected a cell, it either killed the cell by producing numerous copies of itself, or just transformed the cell by incorporating its own genetic material into the genetic material of the infected cell. In the transformed cell the virus was not replicated.

Baltimore and Temin made another major breakthrough in 1970 when they showed the occurrence of a specific enzyme in RNA tumour viruses that could make a DNA copy from RNA. This enzyme was called reverse transcriptase, because it copied genetic information from RNA to DNA as opposed to normal transcription from DNA to RNA. Since this discovery, there has been an explosive growth in our knowledge about the occurrence of genetic material of the type found in RNA tumour viruses in nature. Baltimore, Dulbecco and Temin were awarded the Nobel Prize for Physiology or Medicine in 1975 for their discoveries concerning the interaction between tumour viruses and the genetic material of the cell.

5
Breaking the Code

Even as the mechanism of genetic transmission of traits by DNA was being worked out, efforts were also being made to break the genetic "code" that guided the production of enzyme in the cell. It was known that the DNA double helix contained the complete blueprint of the organism they belonged to; it was also known that the sequence of bases in the DNA molecule made up the code, but the exact manner in which the genetic instructions were coded still remained to be deciphered. It was left to three American biochemists, Marshall Nirenberg, India-born Har Gobind Khorana, and Robert Holley, to independently work out the mechanism.

Even before the genetic code was deciphered, it was realized that it had to be at least a triplet; that is, at least a sequence of three nucleotides in the DNA strand. This is because, if it were a doublet, then, with just four nucleotides to work with, only 16 different combinations were possible. This would not be enough for coding all the 20 amino acids that make up the proteins. But with three nucleotides, the number of possible combinations went up to 64 – more than enough to code for all the 20 known amino acids.

Using the technique developed by Ochoa to synthesize RNA, Nirenberg produced synthetic RNAs made of repeating units of the same nucleic acid, such as uracil (U), adenine (A) and cytosine (C). When he added these synthetic RNAs to cell-free extracts containing no DNA, along with radio-labelled amino acids, he got

Marshall Nirenberg

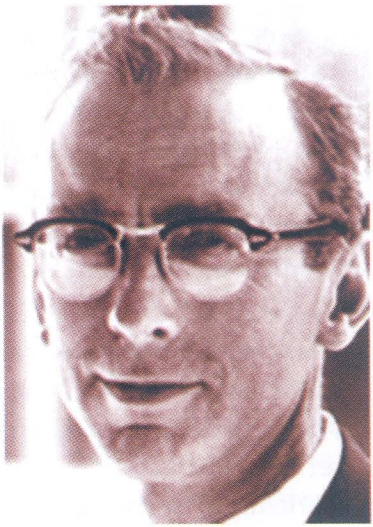
polypeptides containing the same amino acids. For example, RNA containing a chain of only UUU-UUUU..... gave rise to polypeptides composed entirely of the amino acid phenylalanine. Similarly, RNA containing only AAAAAAA..... gave polypeptides containing only lysine, and CCCCCCC..... gave rise to polypeptides containing only proline. Thus it was clear that the nucleotide triplet UUU coded for phenylalanine, AAA coded for lysine, and CCC coded for proline. That is, a triplet of nucleic acids in the DNA chain coded for a single amino acid, as expected.

Nirenberg eventually discovered the codes for virtually all the 20 known amino acids, which are the basic biochemical building blocks. He demonstrated that each possible triplet of four different kinds of nitrogen-containing bases found in DNA (in some viruses, in RNA) ultimately causes the incorporation of a specific amino acid into a cell protein. (Thus a single gene coding for a single enzyme must be composed of several codons.) But DNA also contained sequences that did not code for anything – the so-called "nonsense codons". Nirenberg's teammate Phil Leder subsequently showed that in addition to coding for specific amino acids, DNA also contains sequences that signal the start and the end of the translation process. That is, in addition to words, the genetic code also has punctuation points. In this way Nirenberg showed how the machinery of the cell is used for the translation of the genetic code.

Har Gobind Khorana

Working independently, Khorana confirmed Nirenberg's findings that genetic material is composed of four basic substances and that the way they are linked in large molecules of DNA determines the composition and function of the cell. In course of his research, during which he had systematically devised methods that led to the synthesis of well-defined nucleic acids, Khorana proved that the key combinations come in separate groups of three nucleic acids. He also found that some of the groups prompt a cell to start or stop the production of protein and that some of the amino acids are coded by more than one combination. Khorana's synthetic nucleic acids played a key role in the final solution of the genetic code.

Holley's work related mainly to tRNA. This nucleic acid has the capacity to read off the genetic code from mRNA and to transform it to the corresponding protein in the cell. After years of research with yeast, Holley was able to prepare a tRNA that incorporates the amino acid alanine into

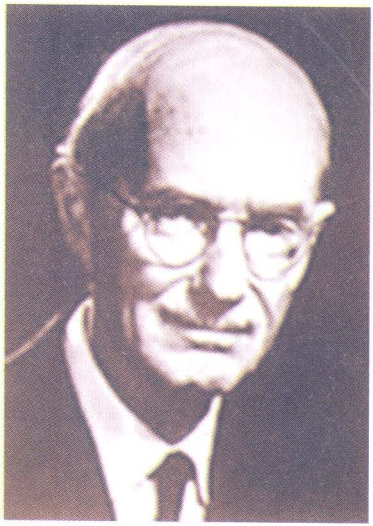


Robert Holley

protein molecules in pure form and, finally, in 1965, to determine its exact chemical structure. He then showed how the tRNA picked up individual amino acids within a cell in a predetermined order and transported them to ribosomes, which then joined them into specific proteins according to the cell's DNA blueprint. For their valuable contribution in deciphering the genetic code, Nirenberg, Khorana and Holley were awarded the Nobel Prize for Physiology or Medicine in 1968.

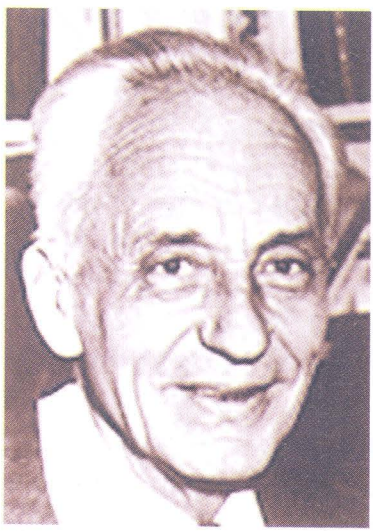
A new direction in the study of gene action came through the work of George Snell, Jean Dausset and Baruj Benacerraf who were able to identify genetic factors responsible for certain immunological reactions. The body's immune system is responsible for the rejection of tissue or organ grafts. From the immunological point of view, every individual is unique. This is true even with regard to the many details of the body's structure. The work of the three scientists demonstrated that the unique character of each individual is determined by genes that regulate the formation of specific protein-carbohydrate complexes, called the histocompatibility antigens, or H antigens, which are found on the cell membrane. H antigens determine the interaction among the multitude of different cells responsible for the body's immunological reactions. Knowledge of H antigens is of great practical importance, for example, in tissue transplantation.

The genetic control of the body's immunological reaction plays a decisive role in the body's defense against infectious agents. People show varying ability to mobilize resistance against infections, and to a large extent an individual's ability to react would seem to be genetically determined. George Snell laid the foundation for our knowledge of the laws that govern the body's ability to distinguish "self" from "non-self". He did this



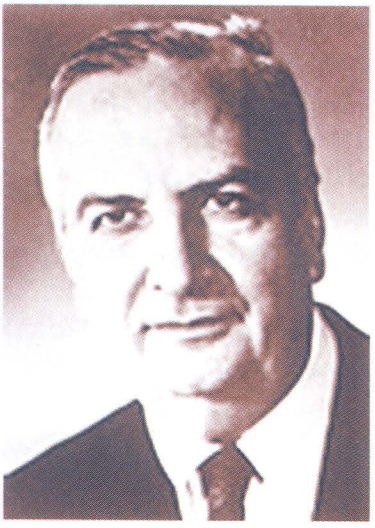
George Snell

by studies on mouse strains that, through repeated sibling mating, were made as genetically identical as monozygous twins. He showed that transplantability was determined by the presence of special structures, which he called histocompatibility antigens, on the surface of the cell. He showed, further, that genes found within a limited area on a specific chromosome (designated as H-genes) controlled the formation of these antigens. This area, which contained a large number of genes, was called the major histocompatibility complex, or MHC. Snell's fundamental discoveries led to the birth of transplantation immunology.



Jean Dausset

The research of Jean Dausset, who identified the human equivalent to the H-genes in mice, was to dramatically blaze the trail to studies of rules for transplantation in humans. Working with antibodies of women who had given birth to several children, he was able to show that one single genetic system, localized on one single chromosome, determined the antigens that gave rise to the

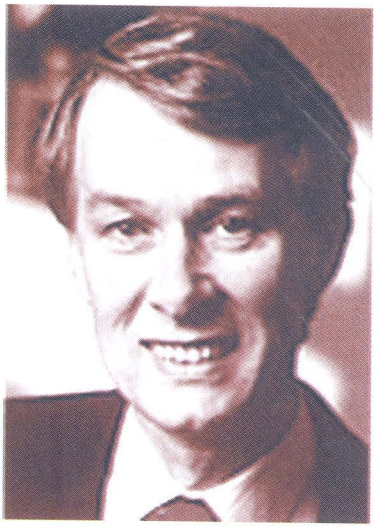


Baruj Benacerraf

antibodies. They came to be called human leukocyte antigens (HLA), and the genes that determined their formation came to be known as HLA genes. Thus, Dausset was able to identify the human equivalent of the H-genes in mice. Dausset's discovery has made it possible to type both donor and recipient in cases of kidney transplantation, thereby increasing the possibilities for a successful transplant.

Benacerraf, through elegant studies with guinea pigs, was able to demonstrate a very important group of genes within the major histocompatibility complex on the chromosome. He showed this region to have several central functions for regulating various immunological reactions in the body, due to which the MHC has come to be referred to as a "super gene." The area of research opened up by Benacerraf made it possible to analyze the background of the varying ability of different individuals to mobilize an immune response to infections. For their discoveries concerning genetically determined structures on the cell surface that regulate immunological reactions, Snell, Dausset and Benacerraf were awarded the Nobel Prize for Physiology or Medicine in 1980.

A new twist was given to the gene story in 1977 when it was discovered that genes could exist as several separated segments. Two scientists – Richard J. Roberts of Cold Spring Harbour Laboratory and Phillip A. Sharp



Richard J. Roberts

of Massachusetts Institute of Technology, made this remarkable discovery independently. Both worked with a common cold-causing virus, called adenovirus, whose genes display important similarities to those in higher organisms. They found that an individual gene could comprise not only one but several DNA segments separated by irrelevant DNA, which came to be known as "split genes." The separate segments of the gene were termed exons and the intervening segments introns. (It is estimated that only 5 per cent of human DNA encodes protein.) Thus a single gene that controls the production of a single enzyme may consist of several exons separated by introns. During transcription, the nucleotide chains folds in such a way as to bring the exons together for mRNA to copy the code.

Phillip A. Sharp

This knowledge has radically changed our view on how the genetic material has developed during the course of evolution. The discovery also led to the prediction of a new genetic

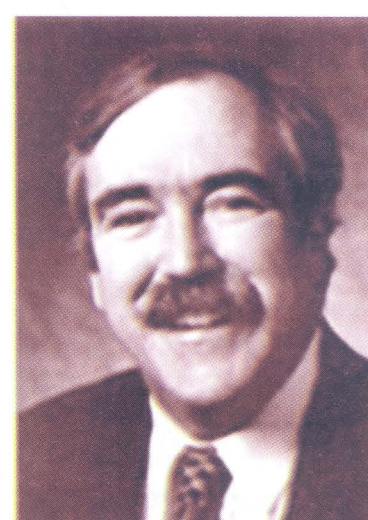
process, namely that of splicing, which is essential for expressing the genetic information. Further, the discovery that genes are often split, threw up the possibility that higher organisms, in addition to undergoing mutations, may utilize another mechanism to speed up evolution – by rearrangement of gene segments to new functional units. This kind of process could drive evolution by rearranging modules with specific functions. It could also help explain the incidence of hereditary diseases like haemophilia or thalassemia, which are believed to be caused due to errors in the splicing process during which DNA is copied through intermediate RNA products. The discovery of split genes has thus been of fundamental importance for not only today's basic research in biology, but also for more medically oriented research concerning the development of hereditary and other diseases. Roberts and Sharp were awarded the Nobel Prize for Physiology or Medicine in 1993 for the discovery of split genes.

The next step in our understanding of gene action came through the work of three developmental biologists who discovered important genetic mechani-sms that control early embryonic development. The three scientists were Edward B. Lewis of California Institute of Technology in Los Angeles, and Christiane Nusslein-Volhard and Eric F. Wieschaus of the European Molecular

Edward B. Lewis

Christiane Nusslein-Volhard

Biology Labora-tory in Heidelberg. All the three worked with the fruit fly and were able to identify and classify genes that are of key importance in determining the body plan and the formation of body segments. They were also able to find out how genes could control the further development of individual body segments into specialized organs. Most of the genes studied by Nusslein-Volhard, Wieschaus and Lewis are known to have important functions during the early development of the human embryo. Mutations in such important genes could be responsible for some of the early, spontaneous abortions that occur in man, and for some of the 40% or so of the congenital malformations that develop due to unknown reasons. Thus the work of the three scientists has had far reaching impact on our understanding of the genetic control of malformations in humans. For their discoveries concerning the genetic control of early embryonic development, Lewis, Nusslein-Volhard and



Eric F. Wieschaus

Wieschaus were jointly awarded the Nobel Prize for Physiology or Medicine in 1995.

The interpretation of the genetic code and the elucidation of its function brought in a revolution in molecular biology, which in the past few decades have led to our present understanding of the detailed mechanism of inheritance. It is now clear how the master molecule DNA, through delicate and carefully orchestrated teamwork with several other molecules, executes the diverse cellular processes that is the hallmark of any living being. It has also led to an understanding of the causes of many diseases in which heredity plays an important role. In fact, these developments formed the very foundation on which the Human Genome Project rested.

6

Splicing Genes

As the genetic mechanism became clear, it also opened up possibilities of tampering with the genes, a technique that came to be known as genetic engineering, and gave birth to a new discipline called biotechnology. Biologists now started thinking of changing the inherited characteristics of organisms in a predetermined way by altering their genetic material, to meet diverse requirements. For example, by introducing genes of wild drought resistant plants into crop plants, crops might be grown with less irrigation. By introducing human genes into bacteria they could be coaxed to produce human hormones and other gene products. The possibilities appeared almost unlimited.

But before such possibilities could be turned into reality, tools for cutting and splicing genes were needed. One discovery that had far-reaching impact on the development of genetic engineering as a tool in molecular biology was that of a special kind of enzymes called restriction enzymes. Restriction enzymes are isolated from bacteria, where they serve as a defense mechanism. They are so called because in bacteria they prevent or "restrict" the incorporation of DNA from foreign sources such as viruses or other strains of bacteria. Restriction enzymes provided the "molecular knives", which the molecular biologist could use to cut strands of DNA into defined fragments. These could then be used to determine the sequence of genes on chromosomes, to identify the regions of DNA that regulate gene function, and to create new combinations of genes.

Werner Arber

Three scientists-Swiss microbiologist Werner Arber, and the American microbiologists Daniel Nathans and Hamilton Smith contributed equally to the elucidation of the function of the new enzyme. Arber and others were only extending the work of Luria, who had observed that bacteriophages not only induce hereditary mutations in their bacterial hosts but also at the same time undergo hereditary mutations themselves. Arber's research was concentrated on the action of protective enzymes present in the bacteria, which modify the DNA of the infecting virus. After discovering the restriction enzymes, Arber postulated that these enzymes bind to DNA at specific sites that contain recurring structural elements made up of specific base-pair sequences.

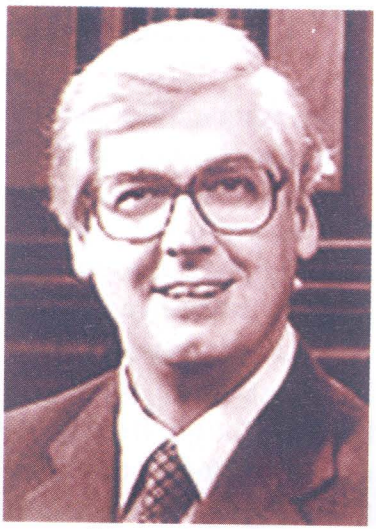
It is now known that each restriction enzyme recognizes a short, specific sequence of nucleotide bases. These regions are called "recognition sequences" and are randomly distributed throughout the DNA. Different bacterial species make restriction enzymes that recognize different nucleotide sequences. When a restriction enzyme recognizes a sequence, it snips through the DNA molecule by catalysing the hydrolysis of the bond between adjacent nucleotides.

Three types of restriction enzymes are known. Type I restriction enzymes split the target DNA molecule at random sites; Type II split the molecule only at the

Daniel Nathans

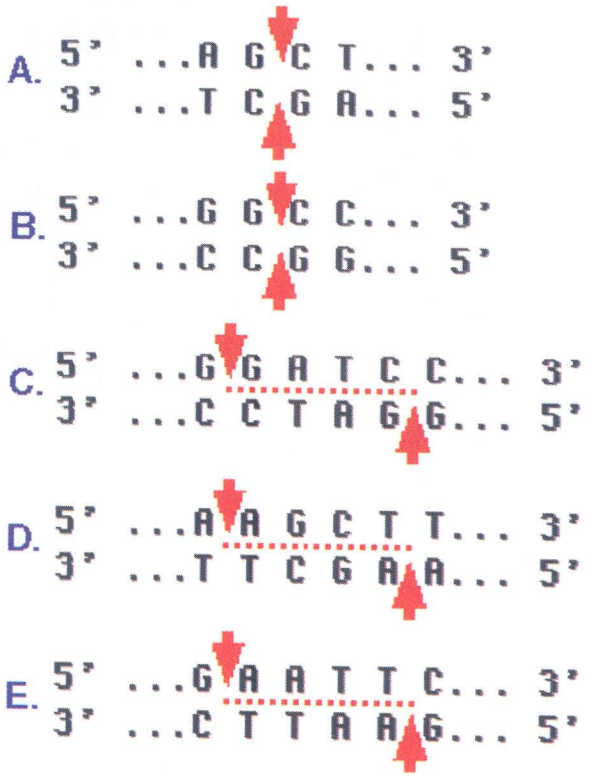
recognition site, while Type III split DNA at a fixed distance from the recognition site. Type II and III restriction enzymes are powerful tools in the elucidation of the sequence of bases in DNA molecules. They play a fundamental role in the field of recombinant DNA technology, or genetic engineering.

Working with the bacterium *Haemophilus influenzae*, that causes influenza, Smith and his colleagues discovered the first of Type II restriction enzymes. These enzymes not only recognize a specific region in a DNA sequence but also cut the DNA at that very site always. This predictable behaviour made type II restriction enzymes valuable tools in the study of DNA structure and in recombinant DNA technology.



Hamilton Smith

Nathan pioneered the application of restriction enzymes to genetics. He used the restriction enzyme isolated by Smith from the bacterium *H. influenzae* to investigate the structure of the DNA of a virus known to produce cancerous tumours in monkey. This



How restriction enzymes cut DNA at specific locations

achievement, the construction of a genetic map of a virus, heralded the first application of restriction enzymes to the problem of identifying the molecular basis of cancer. His work also played an important role in the development of prenatal tests for such genetic diseases as cystic fibrosis and sickle-cell anaemia.

The discovery of restriction enzymes by Arber, Nathans and Smith in 1969 was a significant breakthrough that had far reaching impact on the development of the science of genetics. It led to the founding of a new discipline variously called molecular genetics, genetic engineering, or recombinant DNA

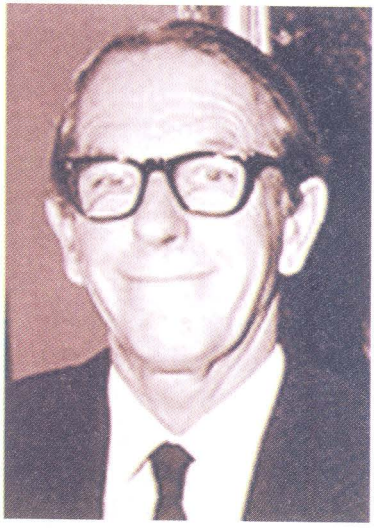
technology. The new biotech tools also opened up new avenues to study the organization and expression of genes in higher animals and to solve basic problems in developmental biology. For their discovery of restriction enzymes and their application to problems of molecular genetics, Arber, Nathan and Smith were awarded the Nobel Prize for Physiology or Medicine in 1978.

Paul Berg

Among those who pioneered techniques of taking a DNA molecule apart and putting the pieces back together were the American biochemist Paul Berg, English biochemist Frederick Sanger and the American molecular biologist Walter Gilbert. It was around 1968 that Berg, who was at Stanford University, turned to the investigation of gene expression in organisms more intricate than bacteria. He and other biochemists had noted that when certain viruses cause tumours, it is the gene of the virus that is expressed in the host body. They decided to study this phenomenon by using techniques that had been developed for chemically manipulating nucleic acids. A simple way to proceed was to cut and install selected pieces – single genes – on a stretch of DNA and study the latter's expression in the new chromosomal environment. While restriction enzymes could do the cutting job, another type of enzyme, called the DNA ligase could be used to stick together two pieces of DNA. Berg used these two enzymes to pioneer the technique now

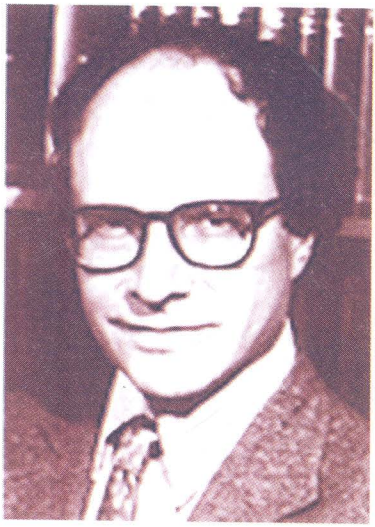
generally called gene splicing, that is, patching together, or recombining, fragments of DNA from different organisms.

Frederick Sanger of the University of Cambridge approached the study of the expression of genes from another angle. He developed techniques to find the sequence of nucleotides in a nucleic acid and correlate it with the proteins that they coded for. Sanger devised a unique method of finding the nucleotide sequence. First, the nucleic acid is cut using restriction enzymes into segments of a few hundred nucleotides, and the segments are separated. Next, each of these segments is used as a template for the synthesis of short, complementary polynucleotides from different mixtures of nucleotides. Analysis of the modified polynucleotides after separation gave the position of specific bases in the original nucleic acid chain. Sanger's method has been used to find the sequence of almost 5,400 nucleotides that make up the DNA of a virus that infects bacteria, and from that, it has been possible to identify nine genes that code for the proteins vital for the functioning of the virus.



Frederick Sanger

The work of Walter Gilbert of Harvard University also concerned the determination of nucleotide sequences in DNA by using special chemical reagents that cleaved the chain at specified locations. In the 1970s, he developed a widely used technique of using gel

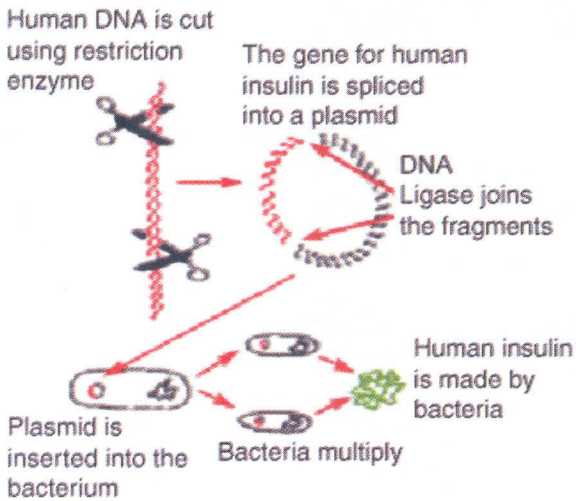


Walter Gilbert

electrophoresis to read the nucleotide sequences of small DNA segments formed by the action of restriction enzymes. For their contribution to the study of nucleic acids, Berg, Sanger and Gilbert were jointly awarded the Nobel Prize for Chemistry in 1980.

The workhorse of the recombinant DNA technology is the common intestinal bacterium *Escherichia coli*. The usual technique of introducing a foreign gene into an organism is to cut the gene to be studied and insert it into the DNA of a virus or a plasmid – entities that can penetrate the cells of animals or bacteria, where the gene can express itself. To insert a foreign gene into *E. coli*, plasmids taken from the bacterium are first cut at selected sites using restriction enzymes. The gene to be inserted is also similarly cut. The two are then joined together using the enzyme DNA ligase, which gives rise to a modified plasmid containing the foreign gene. When the genetically modified plasmid is grown in a culture, all the new bacteria express the inserted foreign gene, producing the gene product coded by it. This is like cutting a printed page in half, inserting a paragraph in the middle and making several photocopies of the page.

This strategy has been widely used in bringing about the expression of mammalian genes in different mammals and in bacteria. Insulin, for example, is a simple protein normally produced by the pancreas. In



Steps in splicing human insulin gene to bacterial plasmid for bacterial production of human insulin

people with diabetes, the pancreas is damaged and cannot produce insulin. Since insulin is vital to the body's processing of glucose, this is a serious problem. Many diabetics, therefore, must inject insulin into their bodies daily. Before the 1980s, insulin for diabetics came from pigs and was very expensive. The animal insulin was slightly different from human insulin and some people were allergic to it.

To create insulin inexpensively, the gene that produces human insulin was inserted into the plasmid of normal *E. coli* bacteria. Once the gene was in place, the normal cellular machinery of the bacteria produced human insulin just like any other enzyme. By culturing large quantities of the modified bacteria and then killing them, the insulin could be extracted, purified and used very inexpensively.

Human insulin, first produced in 1978, was the first pharmaceutical product of recombinant DNA technology. Today many others are being manufactured using bacterial cultures. Human growth hormone is among them. This hormone is used to treat children suffering from dwarfism caused by some defect in the pituitary. Before the days of recombinant DNA technology, human growth hormone was traditionally extracted from human cadavers. But it took hormone extracted from 50 cadavers at great cost, to treat a single child for a year. Human growth hormone manufactured by recombinant DNA technology has changed all that. The product available in the market today is much purer and costs much less than the product extracted from human cadavers.

One of the recent applications of genetic engineering has been in the creation of genetically modified crops and other organisms. The technique has been used to produce bacteria that can gobble up oil spills, plants with an innate resistance to pests and tolerance to herbicides, and food crops with higher nutritional value. Genetic engineering has led to the development of fast-growing and cold-resistant fish, for example, and cheaper, more effective vaccines against livestock diseases as well as livestock feeds that increase the animals' ability to absorb nutrients. In forestry, recombinant DNA technology is being used for enhancing useful traits in plantation trees. The first ever patent for a living organism was awarded in 1980 to the India-born American scientist Ananda Mohan Chakraborty for a genetically modified bacterium that could clean up oil spills.

One of the first genetically modified crops to be released commercially in India is the "Bt cotton", so called because it carries genes from the soil bacterium *Bacillus thuringensis*. In nature this bacterium produces a

Pest-resistant Bt cotton (right) compared with pest-infested ordinary cotton (left)

chemical that kills bollworm, a common pest of cotton plant. When genes from *B. thuringensis* are incorporated in the cotton plant, the plant itself produces the pesticide. So, Bt cotton needs very few pesticide sprays to prevent pest attack. Field trials have shown Bt cotton to give up to 30 per cent higher yield compared to ordinary cotton.

Research is also going on to improve nutritional quality of common food crops using recombinant DNA technology. At Jawaharlal Nehru University in New Delhi, a team led by Prof. Ashis Dutta has already succeeded in producing genetically modified potato that is rich in high-value proteins. The gene in this case was taken from a well-know food plant called amaranthus, which has a high nutritional value because of its well-balanced amino acid composition.

Prof. Dutta's team at JNU has also produced a transgenic oxalate-free tomato variety, by introducing an oxalate-destroying gene from an edible mushroom. The high level of oxalate in common tomato and spinach not only give rise to kidney stones in some people but also makes the crop vulnerable to fungus attack. The

introduction of the oxalate-destroying gene not only makes the transgenic tomato oxalate-free but also significantly increases its shelf life. The next target for introducing the oxalate-destroying gene would be spinach, which, but for its high oxalate content is a highly nutritive leafy vegetable.

Another significant piece of research on genetically modified food has come from the labs of the Swiss plant geneticist Ingo Potrykus and his team, who introduced a carotene-producing gene into rice. Carotene is the precursor of the essential nutrient vitamin A. The genetically modified rice with the carotene gene, also known as "golden rice" because of its golden yellow colour, could be of great value in tackling the widespread deficiency of vitamin A seen among the poor, rice-consuming population of India and many other developing countries. Deficiency of vitamin A is the main cause of malnutrition and blindness in children in India. Nutritional experts hope, carotene-enriched rice may go a long way in supplementing vitamin A in the diet of the poor, who have few other sources of vitamin A. Work has already started on developing an Indian variety of golden rice by introducing the carotene gene into local rice varieties.

At the M.S. Swaminathan Research Foundation in Chennai, a team of young scientists is trying to introduce genes from the mangrove plant, which grows in salt water, into rice genome to make salt-tolerant rice plants. With the threat of rise in sea level as a result of global warming looming large, extensive stretches of rice-growing coastal land are likely to be affected by ingress of seawater. Salt tolerant rice could be an answer to the looming threat of sea level rise.

Not everyone is, however, happy about the tampering with genes. Questions are being raised about the

environmental safety of genetically modified crops and health risks from genetically modified food. Critics of genetically modified crops question the very rationale of genetic modification, which in their assessment could lead to environmental and ecological disaster. Similarly, there are apprehensions about health safety of genetically modified food. But several safeguards have been devised to test for any health or environmental risks from genetically modified organisms before their release into the public domain.

In view of the enormous benefits that genetically modified organisms can bring in the areas of agricultural production, pharmaceuticals and nutrition, it would be unwise to shun them because of unfounded fears. Many of the concerns expressed by critics can be taken care of by careful and extensive pre-release testing and a credible regulation regime. No wonder, despite the misgivings expressed by critics, more and more countries, especially in the Asian region, are going ahead to reap the benefits of the gene revolution that was ushered in by the unraveling of the DNA structure fifty years ago.

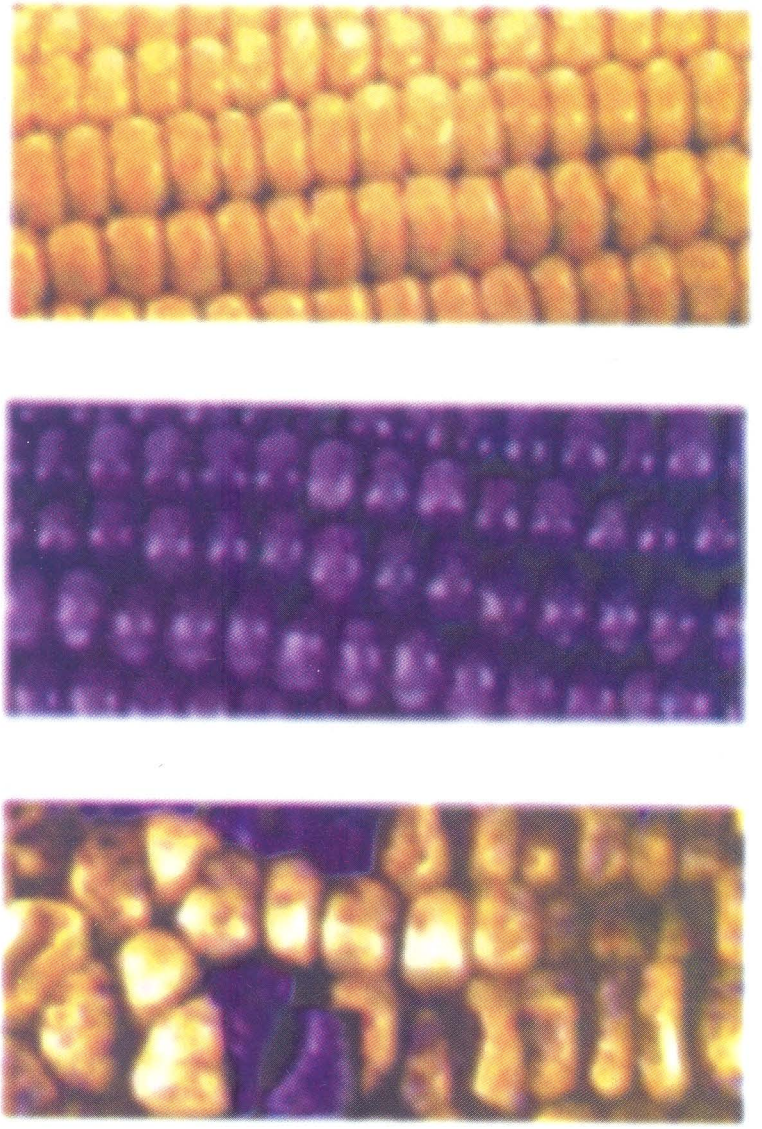
7

Genes that Jump

Long before the DNA structure was unravelled by Watson and Crick and the genetic code was broken, a remarkable discovery about the nature of genes was made by Barbara McClintock, a researcher in Carnegie Institution, Washington, USA. Through a remarkable series of experiments with the common corn plant, *Zea mays*, carried out in the late 1940s, she demonstrated that genes did not necessarily occupy fixed positions on chromosomes, but could jump in and out of chromosomes. These moving genes later came to be known as jumping genes. The concept of moving genes must have seemed incredulous at the time, because DNA was believed to be stable and invariable. It was a remarkable discovery, but since it was made at a time when the real nature of genes were not known, no one really paid any attention to it.

Maize or corn commonly comes with yellow kernels. But in some wild forms of maize grown in Central and South America, the kernels come in purple, brown or red colours. Of course, the colour depends on pigments in the surface layer of the kernel; the colour of the inside of the kernel is always the same – yellowish white. As we have seen with Mendel's experiments with peas, the seed colour in peas depends on the combination of genes received from both parents. So, it was obvious that in the maize plant too the kernel colour should be controlled by the genes of the maize plant.

But in the maize plant something more happens that



Corn grains with uniform colours (top and middle) and with mottled colours (bottom)

is not seen in the pea plant. Sometimes one finds differently coloured kernels on the same cob. The explanation for this is that the corncob is formed from a group of female flowers. If some of these female flowers are fertilized independently by a pollen grain from a different male flower and if the pollen grains do not carry the same genes for kernel pigments, maize cobs with differently coloured kernels would arise. So the appearance of different coloured kernels on the same cob could be explained by the laws of the inheritance stated by Mendel in 1866.

But sometimes, maize kernels show something quite interesting. Rather than being evenly coloured, the kernels show a mottled pattern, with numerous coloured spots or dots on a yellow background, rather than being evenly coloured. This mottling effect defies Mendel's basic principles of genetics because here some individual grains are multicoloured rather than being of a single colour. This was quite puzzling for plant breeders in the 1920s. The problem of mottling in maize was of slight importance from a practical point of view, but it fascinated McClintock because it evidently could not be explained on the basis of Mendelian genetics. She suspected that the dots on the kernels were due to the instability of genes involved in the pigment synthesis. These genes, she thought, underwent mutations during the development of the kernel. If such a mutation were inherited by several generations of daughter cells it would result in a differently coloured spot. This idea received further support when it was found that maize with mottled kernels also had broken chromosomes.

The key to McClintock's success was her development of new techniques of staining that helped her study the chromosomes in greater detail. Using her new techniques, McClintock was able to identify each of

Barbara McClintock

maize's 10 chromosomes and patiently note the location of small bumps, twists, and marks on each of them. Before her work, it was known that maize had 10 chromosomes, but no one had been able to distinguish one from another. McClintock's microscopic observations of the features on each chromosome were so exact that she was not only able to identify each chromosome separately but was also able to observe crossovers between chromosomes through the microscope, and then show that they corresponded precisely to genetic recombination events in the next generation of corn. It was by such careful study of the bumps and twists in the chromosomes and relating them to the kind of mottling produced in the kernel that she found evidence of what she thought was the shifting of the position of genes in DNA. She called these shifting genes "mobile genetic elements".

What did the concept of jumping genes mean for biology? At the time McClintock published her results, genetics had just established the central role of genes on chromosomes in passing the hereditary traits from one generation to another. It had established the rules by which genetic material was sorted, arranged, and recombined, and was beginning to understand how genes were constructed in a chemical sense. To many scientists, McClintock's findings came as a rude shock. They refused to see the maize results as very important. But, in science, the truth rarely remains hidden for long.

Jumping genes were finally isolated from the bacterium *Escherichia coli* in the late 1960s and were further defined as specific, small fragments of DNA which were given the name transposons. In recognition of her path-breaking discovery, McClintock was awarded the Nobel Prize for Physiology or Medicine in 1983.

Once it became clear that jumping genes were a reality, scientific interest in them increased during the 1970s, when it appeared that they assisted in the transfer of bacterial resistance to antibiotics. Today we know that transposons constitute a large fraction of the DNA in some species of plants and animals; among them mice, humans, corn, rice and wheat. It is paradoxical that the discovery of transposable elements lagged so far behind the discovery of the basic laws of genetic transmission. And it is equally curious that even when they were discovered, acceptance of their generality and recognition of their ubiquity came so slowly.

Jumping genes could also provide a plausible explanation for the unique feature of the human immune system. The notion that every cell in the body is genetically identical fails to solve a couple of important genetic problems. One of these is the problem of antibody production. Antibodies, as we know, are proteins produced by the immune system. The most surprising thing about the human immune system is that it is capable of producing an enormous variety of different antibodies when confronted by new antigens. And this is the key to our immune system's ability to fight off disease, even diseases to which it has never been exposed. The number of different antibody proteins that the human immune system can produce is estimated to range from 100,000 to 1,000,000. How does the human body do it?

It is well known that proteins are produced in the cell from mRNAs, which are transcribed from genes in the chromosomes. Therefore, to produce 100,000 antibodies, a human cell must have at least 100,000 antibody genes. But that's impossible because, if that were true, virtually the entire human genome would be taken up just with antibody genes, which is certainly not the case. So the question was, what could be going on? If genes were fixed on chromosomes, there would be no solution to this puzzle.

An important step towards finding a solution to this problem was taken in 1976 by Susumu Tonegawa, a Japanese researcher working at the Massachusetts Institute of Technology in USA. He compared the location of different parts of the antibody gene sequence in embryonic cells with the same parts in cells from an adult mouse. To his surprise, the two parts of the gene, which were together in the adult, were separated in the embryo. What was going on?

The explanation, as Tonegawa showed with a series of elegant experiments, is that the "mature" antibody gene of the adult does not exist in the embryo. Instead, embryonic cells contain hundreds of alternate antibody gene "parts," which must be moved together to assemble the mature, functional gene. And this is what makes it possible for the body to make so many different kinds of antibody proteins. In other words, the embryonic cells have gene parts that can be "assembled" in different combinations to produce a large number of genes that code for antibodies.

In each of the millions of cells of the immune system, bits and pieces are moved from place to place on chromosomes to assemble a functional gene. Since the bits and pieces are selected at random, each cell assembles

a slightly different gene. In effect, in each cell the "deck" of antibody parts is "shuffled" in a different way. The result is that the millions of cells of the immune system contain millions of different antibody-producing genes, preparing the body to face virtually any antigen!

Recent laboratory experiments led by scientists of Johns Hopkins Medical Institutions in USA have revealed that so-called jumping genes create dramatic rearrangement in the human genome when they move from chromosome to chromosome. If the finding holds true in all living organisms, it may help explain the diversity of life on Earth. It is estimated that there are more than 500,000 copies of jumping genes in the human genome, accumulated over the millions of years of human evolution. Most of the changes brought about by transposons were probably disastrous, but occasionally they might have benignly increased genetic variation or even improved the capacity to survive or adapt. Such remodelling probably happened thousands of times during human evolution. Thus jumping genes, which are nothing but stretches of moveable DNA, not only endow organisms with unique traits not explainable by common genetic laws, but also must have played a key role in evolution of life on Earth.

8

Detective DNA

Everyone has a unique set of fingerprints, which is why fingerprinting is so useful for identifying people. For nearly a hundred years, fingerprints have been used to track criminals. They also have been used to identify murder victims and soldiers killed in combat. But fingerprints are not always helpful in catching criminals. People who commit crime often remember to wear gloves or at least to wipe away their fingerprints. Sometimes, even when police find a print, they can only compare it to the ones they have on file. If the culprit has never been arrested before, police won't find a match. Sometimes, fingerprints are not easily available as evidence, especially from scenes of a crime such as rape or murder.

Today, the ubiquitous DNA has turned out to be a more reliable identifier of individuals than any available so far. Unlike conventional fingerprints, where the pattern of ridges in the skin is matched for identification, in DNA fingerprinting (also known as DNA typing) the unique pattern formed by the nucleotide sequence of certain stretches in the DNA of an individual is matched for identification. These patterns do not, however, give an individual "fingerprint," but they are able to determine whether two DNA samples are from the same person, related people, or non-related people.

Scientists use a small number of sequences of DNA that are known to vary among individuals a great deal, and analyse those to get a certain probability of a match. The method is so reliable that the probability of any two DNA fingerprints being the same by pure chance is less

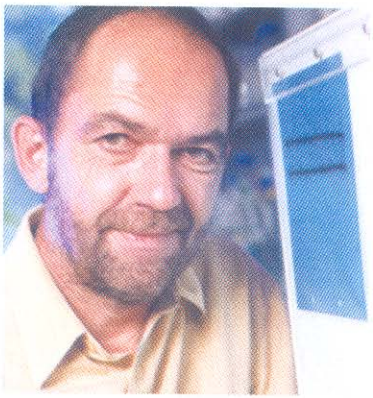
than one in several thousand million. Compared to conventional fingerprints, one of the great advantages of DNA typing is that there are so many ways to get a "print". One can use hair, blood, saliva, semen, skin, and even nail clippings, because they all are made up of cells containing DNA.

Every one of us humans have 23 pairs of chromosomes containing the DNA blueprint that encodes all the materials needed to make up our body as well as the instructions for how to run it. One member of each chromosomal pair comes from the mother, and the other is contributed by the father. Every cell in our body contains a copy of this DNA. While the larger part of DNA does not differ from one person to another, some 3 million base pairs of DNA, or only about 0.10 per cent of the total DNA, vary from person to person. The key to DNA fingerprinting lies in the uniqueness of those regions of DNA that do differ from person to person.

Amazingly, despite our complexity, genes make up only 5 percent of the human genome and individual variations within genes are not very useful for DNA fingerprinting. The other 95 percent of our genetic makeup does not code for any protein. Because of this, these non-coding sequences used to be called "junk DNA." But, it turns out that these regions do actually have important functions. It has been found that they contain repeated sequences of base pairs. These sequences, called "variable number tandem repeats", or VNTRs, may contain anywhere from 20 to 100 base pairs. In such stretches, a particular sequence can be repeated anywhere from one to 30 times in a row. A given person's VNTRs come from the genetic information donated by his or her parents. So, he or she could have VNTRs inherited from his or her mother or father, or a combination of them, but never a VNTR that has not

come from either of his or her parents.

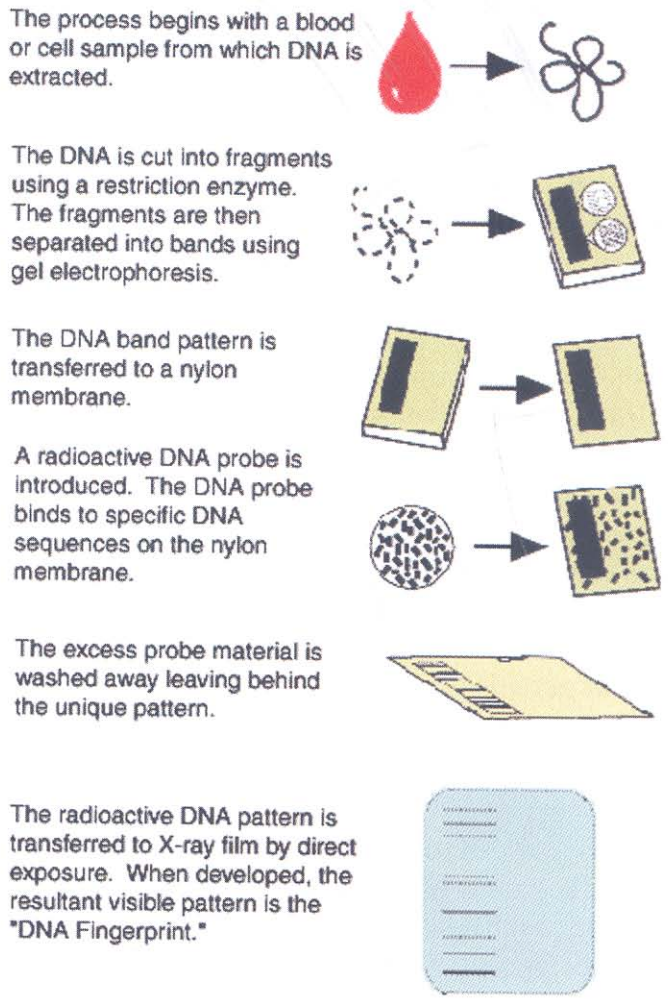
The DNA fingerprinting technique was developed in 1984 by the British geneticist Alec Jeffreys, after he noticed the existence of certain repeating sequences of DNA that were later termed as VNTRs. He also found that each individual has a unique pattern of these VNTRs, the only exception being identical twins. So, if the pattern of these repeat units could be deciphered, the identity of the individual from whom the sample had come could be established unequivocally.



Alec Jeffreys

DNA fingerprinting basically involves translating the VNTRs of suspects into visual records, which can be matched with samples obtained from known sources, including the scene of crime or the body of a victim. DNA is first extracted from a sample and cut into segments of various lengths using restriction enzymes. The resulting fragments are tested for what is known as "restriction fragment length polymorphism", or RFLP. It focusses on segments of VNTR fragments that contain certain sequences of repeated DNA bases, which vary widely from person to person.

To test for RFLP, the DNA fragments of varying lengths are sorted according to size by a technique known as gel electrophoresis in which the mixture of fragments is placed on a gel and subjected to an electric current. When an electric current is passed, the fragments move toward the positive pole at speeds depending on their size. The shorter the fragment faster they move toward



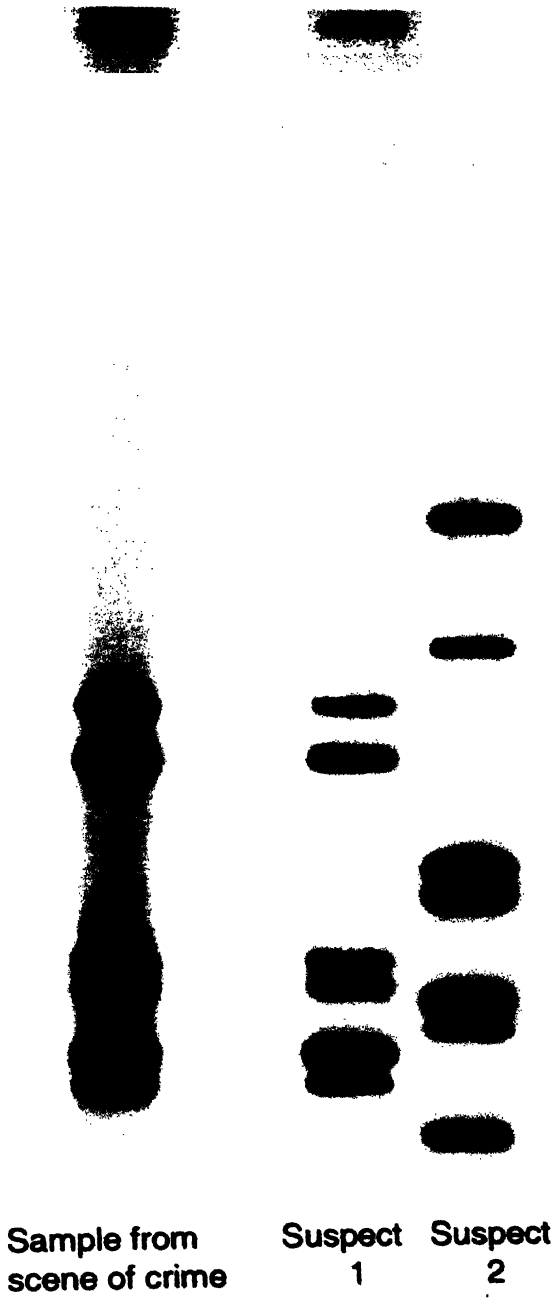
Steps in making DNA fingerprints

the positive pole. At the end, bands of DNA fragments, arranged according to size, are obtained. The sorted, double-stranded DNA fragments are then subjected to heat or chemically treated to break up the double strands of DNA into single strands. The band pattern of single-stranded DNA obtained by electrophoresis is then

transferred (blotted) to a nylon or nitrocellulose membrane. This is known as Southern blotting (after the British biologist Edward M. Southern, who developed it).

When a single-stranded DNA probe, tagged with a radioactive isotope, is then applied to the Southern blot, the specially designed probe binds to specific VNTR segments of single-stranded DNA on the nylon or nitrocellulose membrane. After the excess probe material is washed away, a unique DNA pattern with part of it containing radioactive isotopes is left behind. If a sheet of X-ray film is now placed in contact with the Southern blot, only the areas where the radioactive probe binds show up on the film. The resulting pattern of dark bands on the developed X-ray film is the DNA fingerprint, which allows researchers to identify, in a particular person's DNA, the occurrence and frequency of the particular genetic pattern contained in the probe. If only a small amount of DNA is available for fingerprinting, a technique called polymerase chain reaction, or PCR is used to multiply DNA.

PCR is a technique used to make numerous copies of a specific segment of DNA quickly and accurately. The technique enables investigators to obtain the large quantities of DNA that are required for various experiments and forensic analysis. The technique was developed in 1983 by the American biochemist Kary B. Mullis. It is a three-step process, based on the natural processes a cell uses to replicate a new DNA strand. Only a few biological ingredients are needed for PCR. Before the development of PCR, the methods used to amplify, or generate copies of DNA fragments were laborious and time-consuming. In contrast, PCR is a fully automated technique. A machine designed to carry out PCR reactions can complete many rounds of replication,



DNA fingerprints showing match with suspect 1

producing thousands of millions of copies of a DNA fragment, in only a few hours.

In India, pioneering work in developing DNA probes for DNA fingerprinting has been done by Lalji Singh at the Centre for Cellular and Molecular Biology, Hyderabad. He developed the probes using DNA sequences isolated from the poisonous Indian snake called the banded krait.

DNA fingerprinting has today become an integral part of the criminal justice system; prosecutors in many countries are using it increasingly to identify and convict criminals. Although its reliability was disputed on scientific grounds for several years, by the mid-1990s DNA fingerprinting had been sufficiently refined to gain wide acceptance in both the scientific community and the courts. Forty-three states in USA now maintain DNA databases of convicted criminals, a resource that assists in identifying the perpetrators of new offences.

At the same time, DNA has also acted as an instrument of innocence. Since 1991, lawyers of a group called the "Innocence Project" in New York, USA have been using DNA typing as evidence to help exonerate people they believe were wrongly convicted. Largely because of the project, by 2001 more than 100 prisoners, many of them on death row, were acquitted on the basis of DNA typing.

DNA has thus unexpectedly spotlighted a need for reform of criminal justice. Its use in exonerating convicts has exposed serious flaws in the conventional system of prosecution, especially in cases in which capital punishment is awarded. The shaky reliability of eyewitnesses and the inadequacy of conventional forensic data in establishing identity have been exposed.

DNA typing is also being used by plaintiffs seeking to

prove paternity, by forensic scientists to identify murderers and rapists, and by analysts attempting to identify victims of disasters. In the famous "Tanduri murder" case in India some years ago, the identity of the victim could be established from half-burnt bone fragments only by using DNA fingerprinting. In the 1990s, DNA test was used to link remains to passengers who died in the crashes of TWA flight 800 and the Swiss Air flight off Nova Scotia. It also helped determine who were buried in the mass graves in Bosnia. After the attack on the World Trade Centre in New York on September 11, 2001, expectations ran high that DNA would help identify the remains of victims at the attack site. Indeed, by March 2002, DNA typing of the remains at the WTC site had led to the identification of about 200 of the victims.

DNA typing has also revolutionised anthropology. Using DNA to trace human lineages, anthropologists have found that the deeper but still intimate ties between Europeans, Asians, the peoples of the Americas and Oceania, all now seem to lead back 100,000 or 200,000 years to a single woman in Africa. DNA studies have also established a common link among the all members of the animal kingdom, especially the link between humans and the apes, as postulated by Charles Darwin almost a century-and-a-half ago. DNA thus makes nonsense of the old ideas of human superiority and of race – those notions of purity and separateness so dear to racists.

9

Mapping the Human Genome

A major landmark in DNA research was reached in 2000 when the first draft of the human genome was made public. Genome is the complete set of genes present in an organism. The human genome contains almost 3.1 billion sub-units of DNA – the chemical "letters" – packed in the 23 pairs of chromosomes that make up the recipe of human life. The importance of the success with the human genome stems from the fact that it holds the key to almost everything that defines a human being, including the physical traits, habits and more

The 23 pairs of human chromosomes, which have been completely mapped under the Human Genome Project. Each parent contributes one chromosome to each pair

importantly, proneness to certain diseases. Many scientists believe that armed with the genomic data they can better understand the functions of genes and correlate genetic abnormalities with specific diseases. This could enable doctors to find out whether an individual is genetically predisposed to develop certain diseases later in life and, may be, to recommend preventive measures, or even corrective action using newly developed drugs or treatments.

The Human Genome Project was started formally in 1990, originally as a 15-year project coordinated by the US Department of Energy and the National Institutes of Health. But thanks to rapid advancements in gene sequencing and computer technology, the first draft of the entire human genome was ready by 2000; that is, five years before the target date. The primary objectives of the project were:

- To identify all the approximate 30,000 genes in human DNA;
- To determine the sequences of the 3 billion chemical base pairs that make up human DNA;
- To store this information in databases;
- To improve tools for data analysis;
- To address the ethical, legal, and social issues that may arise from the project.

Long before there was a formal Human Genome Project, however, the US Department of Energy and its predecessor agencies had been interested in developing sensitive methods to detect changes to genetic materials induced by ionising radiation such as X-rays and nuclear radiation and to understand the related health effects. It has been known for some time that the genetic-information-containing DNA is the part of a cell that is

the most sensitive to the effects of radiation and other pollutants, even in low doses. As new technologies for understanding and working with DNA were developed in the 1980s, the idea arose to sequence the entire human genome systematically, and the Human Genome Project was conceived. It was recognized early on that once this project was completed, it would furnish a comprehensive reference source that others could build on without having to repeat the research from scratch.

When the first draft of was released in June 2000, however, it sprung many surprises. Scientists who were trying to count the number of genes in the human genome could never expect what was in store for them. So, when the results finally came, they were all baffled. Contrary to expectations, humans were found to carry only about 30,000 genes compared to the 19,000-odd genes carried by the lowly roundworm. This was less than a third of the estimated 100,000 genes that the human genome was earlier believed to contain.

The findings were really puzzling, as humans are known to be far more advanced on the evolutionary scale and more capable intellectually than a lower invertebrate like the roundworm. For example, the roundworm is a little tube-like creature with a body made up of only 960 cells, of which 300 are nerve cells that make up its "brain". In comparison, humans have 100,000 billion cells in their body, including 100 billion brain cells. How could the small number of genes account for the far greater complexity of the human organism?

The researchers have also come out with a possible answer. According to them, humans are very thrifty with their genes and are able to do more with what they have than other species. For instance, instead of producing only one protein per gene, as believed earlier, the average human gene has been found to produce three different

proteins. According to one of the researchers, with 30,000 genes, each directly interacting with four or five on average, the human genome is not significantly more complex than a modern jet plane, which contains more than 200,000 unique part, each of them interacting with three or four others on average.

Another surprising discovery was that in the 75 percent of repetitive DNA sequences, which are known as junk DNA and were considered to be useless, there are sequences that are still active and may be coding for proteins. Besides, it is possible to date groups of so-called "repeats" in junk DNA to when in the evolutionary process they were "born" and to follow their fates in different regions of the genome and in different species. Based on such "DNA dating", scientists can trace the genealogy of the human race. The research teams also found evidence of how DNA changes over time, causing organisms to evolve.

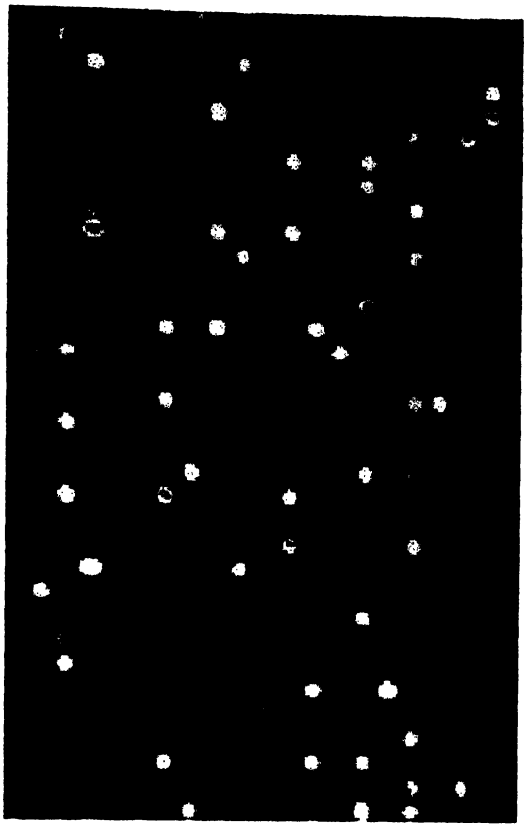
The goal of the Human Genome Project is ultimately to produce a completely finished gene map, with no gaps and 99.99 percent accuracy. Despite the many imponderables still remaining to be sorted out, the mapping itself marks a significant step forward in our understanding of the human system, and may be, towards finding cures for many debilitating afflictions of mankind.

It may quite some time before the newfound knowledge gained from the mapping of the human genome could be put into practice, although some headway has already been made in understanding a few genes linked to diseases. This is because, the human genome is known to be mostly junk DNA that do not code for anything. Besides, except for a few disorders, specific genes for most diseases are yet to be identified –

a task that may take several years to complete. And finally, even if specific genes were identified, they would not mean actual incidence of the disease, because, as it is now known, a single gene may not be responsible for a single disease. Besides, expression of many genes is influenced by environmental and other factors.

Nonetheless, technology and resources promoted by the Human Genome Project are starting to have profound impacts on biomedical research and promise to revolutionise the wider spectrum of biological research and clinical medicine. Among the novel technologies that have come up are the gene sequencers, machines that have automated the biochemical tests necessary to identify genetic sequences using gene probes. These computerised machines can sequence DNA thousands of times faster than is possible manually. Without these fantastic machines the mapping of the human genome could not have been completed five years ahead of target.

Another significant innovation engendered by the Human Genome Project is the so-called DNA chip (also known as DNA microarray), which makes use of the complementary pairing of DNA strands to identify thousands of genes simultaneously, using probes made of known segments of single-stranded DNA. In designing a chip, strands of identified DNA are made and purified using conventional techniques such as polymerase chain reaction and biochemical synthesis. Short lengths of these single-stranded DNA fragments (5 to 25 base pairs) with known base sequences are carefully arranged on a wafer of glass or silicon. Using the technique of photolithography, as many as 300,000 distinct sequences representing over 6,000 genes can be put on a single 1.3 cm × 1.3 cm DNA chip. When such a chip is brought into contact with a mixture of foreign DNA or RNA fragments, only those strands in the



A typical DNA chip showing positions of identified genes as fluorescent spots

mixture that have nucleotide sequences complementary to those on the chip bind to the single-stranded DNA fragments of known sequence on the chip. Special chemicals that bond to successfully paired nucleotide sequences in the chip help researchers identify results. Since such chemicals usually show fluorescence, it is possible to identify specific genes and measure levels of their expression by scanning the surface of the chip with a laser scanning device and measuring the intensity of fluorescence at each position in the chip. Besides hastening the mapping of the human genome, the DNA chip technology is also having a significant impact on

genomics study. Many fields, including drug discovery and toxicological research, are likely to benefit from the use of this technology.

Increasingly detailed genome maps available today have helped researchers seeking genes associated with dozens of genetic conditions, including muscle dystrophy, inherited colon cancer, Alzheimer's disease, and familial breast cancer. Researchers in the field expect that DNA chips will enable clinicians—and in some cases even patients themselves—to quickly and inexpensively detect the presence of a whole array of genetically based diseases and conditions, including AIDS, Alzheimer's disease, cystic fibrosis, and some forms of cancer, as blood sugar tests are available today.

On the horizon is a new era of molecular medicine characterized less by treating symptoms and more by looking to the most fundamental causes of disease. Rapid and more specific diagnostic tests will make possible early treatment of several ailments. Medical researchers also will be able to devise novel therapeutic regimens based on new classes of drugs, and immunotherapy techniques. The new knowledge may also make it possible augmentation or even replacement of defective genes through gene therapy.

The availability of a virtually complete list of human gene products may give us a vast repertoire of potential new drugs. From 500 or so drugs in 2000, at least six times this number will have been identified, tested, and commercialized in 2020. Most probably, all will be manufactured by recombinant DNA technology so they will be "reagent-grade pure," just as human insulin and growth hormone are today.

Already the Human Genome Project has created the field of genomics – the study of genetic material on a

large scale. The medical industry is building upon the knowledge, resources, and technologies emanating from the project to learn more about genetic contributions to human health. As a result of this expansion of genomics into human health applications, the field of genomic medicine has been born. Genetics is playing an increasingly important role in the diagnosis, monitoring, and treatment of diseases.

All diseases have a genetic component, whether inherited or resulting from the body's response to environmental stresses like viruses or toxins. As a result of the success of the Human Genome Project, it will now be possible for researchers to pinpoint errors in genes – the smallest units of heredity – that cause or contribute to disease. The ultimate goal is to use this information to develop new ways to treat, cure, or even prevent the thousands of diseases that afflict humankind.

In future, understanding genomics is expected to help us understand human evolution and the common biology we share with all of life. Comparative genomics between humans and other organisms such as mice already has led to discovery of similar genes associated with diseases and traits. Further comparative studies will help determine the yet-unknown function of thousands of other genes.

But the road from gene identification to effective treatments is long and fraught with challenges. In the meantime, biotechnology companies are racing ahead with commercialization by designing diagnostic tests to detect errant genes in people suspected of having particular diseases or at risk for developing them. An increasing number of gene tests are becoming available commercially, although the scientific community continues to debate the best way to deliver them to the

public and medical communities that are often unaware of their scientific and social implications. While some of these tests have greatly improved and even saved lives, scientists remain unsure of how to interpret many of them.

The Human Genome Project has a flip side too. Misuse of the newfound knowledge cannot be ruled out altogether. For example, patients taking the tests face significant risks of jeopardizing their employment or insurance status unless absolute confidentiality of such knowledge is made mandatory. And because genetic information is shared, these risks can extend beyond them to their family members as well. Another danger is that the new knowledge may give rise to genetic determinism – the belief that human characteristics such as intelligence or criminal behaviour are a function of genes – which, in turn may be used to perpetrate racial or ethnic inequalities.

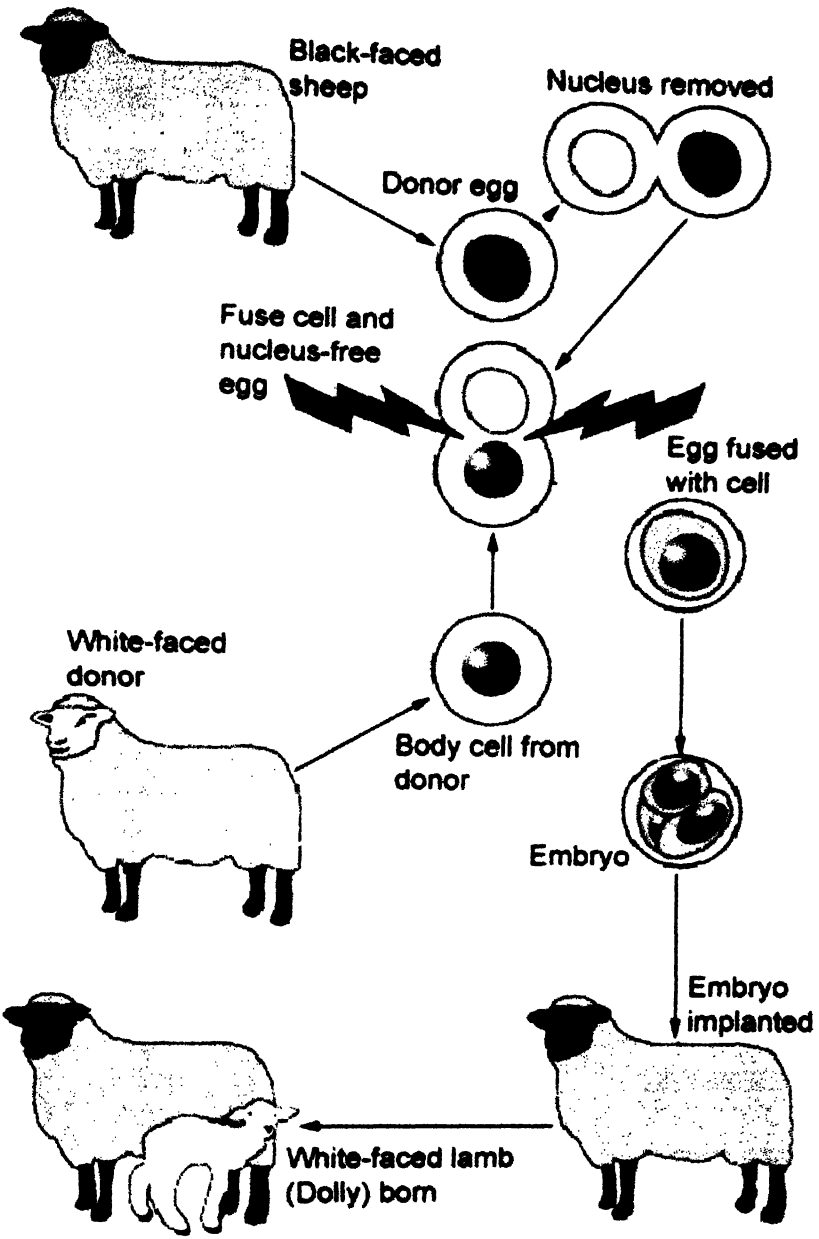
The availability of a catalogue of tens of thousands of genes has also give rise to a new discipline of large-scale studies of protein structure and function, which has come to be known as proteomics. This has become necessary because, although DNA transmits the genetic information, in the cell it basically has a passive role. It is the proteins encoded by DNA that actually carry out the myriad cellular reactions that ultimately constitute "life." Scientists are therefore trying to find out what all the proteins made by these genes actually do, and they are using a novel method. Since mice breed quickly and share about 99 per cent of their genes with humans, scientists are trying to learn about the function of the numerous proteins by studying how mutant genes affect protein function in the mice. Such studies, they hope, will help us make sense of the tens of thousands of genes discovered by the Human Genome Project.

The recent success in mapping the human genome is only the first small step in our understanding the book of life. The great leap is still awaited. The alphabets have been read, but most of the words and the sentences are yet to be deciphered and their meanings understood. Until that is done, the book of life will remain just a magnificent computer database. Nonetheless, the successful mapping of the genome itself is indicative of tremendous stride the ubiquitous DNA has made since Watson and Crick unveiled its double-helix structure fifty years ago. It not only has led to the unravelling of the book of life but also given rise to a multi-billion dollar global industry that may hold the key to the future of mankind.

10
Making Copies of Life

It has long been known that that every cell in an with exception of the sex cells, contains the entire blueprint necessary for the development of an ind But soon after the fertilised egg starts its life as a ε embryo, most of these cells change into speciali types, such as blood, hair, skin, bone, and nerve c these specialised cells, regulatory genes typica down all "unnecessary" genes so that nerve cells functions related to nerves and muscle cells cor only to muscles. Thus it was believed that onc become specialised for a particular function, a c never be reprogrammed to behave as an embry to produce a whole animal. That is, making a cl grown up animal was thought to be impossible. birth of the cloned sheep "Dolly" in 1997 changec

The technique of cloning has been known fc time. The technique used by botanists and floric to multiply plants by tissue culture using only a from fully-grown plants is a form of cloning, ar produced by this technique are all genetically i It goes without saying that the technique allo multiplication of elite varieties of economically ir plants and flowers, not possible by conv cultivation methods. But all past efforts to clone especially from adult cells, had ended in failure, some successes were reported in cloning embry This was not surprising, because embryonic known to be totipotent; that is, they have the ca



Steps used to clone Dolly

differentiate and grow into a full organism and were amenable to cloning. But once differentiated, or transformed to perform specialised functions, cells were believed to lose their totipotency; that is, lose their ability to be grown into a full organism. So, it had been the generally accepted dogma that adult mammalian cells were incapable of differentiating into the various cell types necessary to make a complete and viable mammal, till Dolly proved it otherwise.

What Ian Wilmut and his colleagues at the Roslin

Dolly (left) with her surrogate mother

Institute in Edinburgh, Scotland, did was to transfer the entire nucleus, containing the full complement of DNA, from the body (udder) cell of a grown up female white-faced sheep into the egg cell of a black-faced sheep from which the nucleus had been removed. This was done by putting the body cell and the nucleus-free egg cell together and exposing them to an electric current as a result of which the two fused, forming an egg cell carrying the inserted DNA from the donor. It was now like a fertilised cell, although the entire complement of genes in this case came from a single animal and not from two parents, as happens with sexual reproduction. The fused cell soon started to divide like any fertilised egg, producing an embryo, which was implanted in the uterus of a third, black-faced female sheep to develop to full term. Dolly was born a few months later. Since it carried DNA identical with the DNA of the white-faced sheep from which the body cell was taken, Dolly was an identical copy, or clone, of that sheep. It was the first mammal ever to be cloned.

Since Dolly, several university laboratories and companies have used various modifications of the nuclear transfer technique to produce cloned mammals, including cows, pigs, monkeys, mice and most recently, a gaur – an endangered species of wild ox found in India.

When the news of Dolly's birth was announced in 1997, it immediately raised the spectre of creating human clones and its possible consequences. Even at that time the idea was considered unethical and dangerous, and restrictions were immediately imposed in many countries on experiments with human cloning. Leading experts in reproductive biology warned that human cloning would inevitably lead to babies with genetic defects, or that such babies would die soon after birth. But there were protagonists who didn't see any danger in such an

endeavour and there always remained the possibility of enterprising researchers getting around the restrictions to have their way. In December 2002, media reported the birth of the first cloned baby, although the claim has not been scientifically proved.

Although the idea of artificially cloning humans is reprehensible, there is nothing really new about human clones. Naturally occurring genetic clones, or individuals genetically identical to one another had long been recognised in the form of monozygotic (identical) twins, triplets, and so on. Such clones are derived, as their scientific name indicates, from a single zygote, or fertilized egg, which divides within the womb before implantation, leading to the simultaneous development more than one embryo. But creating human clones artificially has never been tried before. In fact, as mentioned above, although clones had been generated previously in the laboratory from embryonic cells of lower animals such as frogs, decades of attempts to clone mammals from adult cells had met with repeated failure, till Dolly was born.

However, the success of Dolly did not mean that all hurdles encountered earlier in attempts to clone adult mammalian cells had been overcome. The apprehensions about human cloning still remain valid and the risks of cloning a human still remain too high to be taken lightly. The birth of Dolly itself is illustrative of the risks involved; it was just one success out of 227 attempts. In experiments with other mammals such as mice, pigs and goats, too, the fatality rate was found to be very high; from 20 to 40% of the cloned animals died soon after birth or suffered from serious genetic defects. Even after a normal birth, Dolly did not survive for long; it was afflicted with old age problems like arthritis and lung congestion at a "young" age of only six years, compared to a normal

lifespan of 11 to 12 years for healthy sheep. Certainly, this state of affairs cannot be acceptable for human experimentation.

Even if these hurdles were overcome, human cloning still is not worth the effort. There are several important issues that need to be addressed before human cloning can be considered viable. Among the scientific issues, there is the question of biological relevance and medical risks. Unlike in animals, it is well known that genes play a very small part in human intelligence and mental capability – qualities that distinguishes humans from other animal species. What is important here is that, a cloned human offspring can never become an exact replica of an individual from which it is cloned because what a newborn human baby ultimately grows up into depends not so much on its genetic make-up as on its training and upbringing. So, even if an Einstein or a Jagadish Chandra Bose were to be cloned there is no guarantee that the clones would grow up to be as intelligent as the great scientists, although they would be exact look-alikes. Thus, the very advantages of cloning, namely creation of identical copies with desirable traits, as applied to cloning plants or animals, would not apply here.

Also there is no biological relevance of the penchant for having a "genetically related" offspring because it does not really offer any real advantage to the parents, apart from the mental satisfaction of having an offspring of one's own. But if a childless couple wants to have a child of their own they can go for any of the several safe and ethically acceptable alternatives available. These include artificial insemination, in-vitro fertilization, and even adoption. If one does not suffer from a mental block, an adopted child can be cared for and loved as much as one's genetically related child. Hundreds of thousands

of childless couples around the world, who have tasted the joys of parenthood by adopting these alternatives, are a shining testimony to this.

But of greater concern are the medical risks involved in human cloning. As mentioned above, the survival rates in animal experiments are too low to be acceptable to be used for humans. There is another catch. To create Dolly, the nucleus from the udder cell of an adult ewe was fused with an ovum of another ewe from which the nucleus had been removed. Thus the entire complement of genes of Dolly had come from a single female and there was no male contribution, as in normal conception. Thus the mixing of genes, which is the hallmark of sexual reproduction, was prevented, thereby striking at the very root the creation of genetic diversity in the population the implications of which in the long run could be disastrous.

Further, it is known that adult body cells are derived by the repeated division (copying) of cells, which renew almost all the cells of the body after periodic intervals. It is a renewal process that goes on throughout our life. But every time a cell is copied it is actually a "copy of a copy". Every time a cell divides and makes a copy of itself, minute errors are introduced in the copied genetic material (DNA), which can go on only up to a limit. In fact, this accumulation of errors in the genetic material is at the root of the process we call aging, and all living cells are genetically pre-programmed to divide only a certain number of times that decides the lifespan of a species. So the question would naturally arise about the prudence of using the nuclear DNA from an adult human cell for cloning, as was done with Dolly, which died a premature death.

At the ethical level, the issues are different. An offspring born through the technique of cloning will carry

genes from only the father or only the mother from whose cell the nucleus has been taken, rather than from both parents, as in a normal offspring. As a result, the offspring born will be in reality the genetically identical brother of the "father" if male, or genetically identical sister of the "mother" if female, depending on from whom the nuclear DNA is taken. Such a situation would not only be absurd but could also have serious repercussions on the mental development of the child born this way and on the society as a whole. It will also have an impact on gender relationships. For instance, by using the technique a woman could decide to have a girl child without the participation of the husband, but she could never produce a male child of her own. At the same time, male participation would produce only male offspring but no female offspring. Such a state of affairs could have serious social implications.

However, there is also a brighter side of human cloning. While cloning entire humans could be dangerous and undesirable, some supporters of human cloning look at it from a different angle. According to them, cloning offers a viable technique of growing human tissue and organs for transplantation. Organ transplants to replace a diseased or damaged organ of the human body is quite commonplace today. But the problem is that the right donor organ with matching tissue type is not always available. Now scientists in the US and India have succeeded in cultivating a special type of cells called stem cells in the laboratory and coax them to grow into any desired type of tissue. Stem cells are primitive cells that have two remarkable properties. First, these cells are, in principle, immortal. Whereas most cells divide a finite number of times and perish, stem cells can be cultured to divide indefinitely, which makes them excellent objects for manipulation by researchers. Secondly, stem cells

are pluripotent; that is, they can turn into all cell types such as muscle, kidney, liver, blood and even nerve cells.

Stem cells are usually found in very early embryonic stage called blastocyst and early work on stem cells was done only with embryonic stem cells, which led to a lot of controversy, as killing embryos was considered unethical. But now permission has been granted for research with existing lines of embryonic stem cells maintained at various centres around the world including some centres in India, opening up a new vista of research on regenerative medicine.

Meanwhile, scientists have found other sources of stem cells that can be tapped for tissue or organ regeneration without raising any ethical issue. Among them are bone marrow, and placental blood. Researchers in the US and India have already succeeded in regenerating heart muscles, fallopian tubes, liver cells and even nerve cells in animal experiments using stem cells. Regeneration of heart muscle and nerve cells could open up possibilities for treating patients of heart attack and stroke. Some researchers predict that such research will progress so fast that human trials of prospective treatments might begin within a decade. If that happens, it may become possible to grow anything from heart muscle to bone marrow and brain tissue in the laboratory for transplantation.

Thus cloning is a powerful technique that can play an important role in animal husbandry for multiplying elite animals. It may also greatly simplify the otherwise cumbersome manipulation of domestic livestock currently required for genetic improvements in resistance to disease. It may also facilitate the production of life-saving pharmaceuticals for human use, for example, the production of human insulin in buffalo or cow milk. In

addition, cloning of selected animals can also help in saving endangered species otherwise doomed to extinction, as has been done with the gaur recently. At the same time, while cloning of human stem cells could help develop new techniques of regenerative medicine, cloning of entire human beings would not provide any additional advantage over the other alternatives presently available to be worth the risks involved.

Milestones in DNA Research

1866 Gregor Mendel publishes his work on pea plants, describing the nature of heredity.

1869 Johann F. Miescher discovers DNA, which he calls "nuclein".

1909 Phoebus Levene discovers that the D in DNA is deoxyribose.

1915 Thomas H. Morgan establishes the link between chromosomes and heredity.

1944 Oswald T. Avery, Colin McLeod and Maclyn McCarty show that genetic information is stored in DNA.

1948 Barbara McClintock predicts the existence of "jumping genes."

1950 Erwin Chargaff discovers that in all species there is a 1 to 1 ratio of the bases A to T and between C to G.

1953 James D. Watson and Francis Crick discover the double-helix structure of DNA.

1961 Francois Jacob and Jacques Monod identify the role of messenger RNA and regulator genes.

1966 Marshall Nirenberg, Har Gobind Khorana and Robert W. Holley crack the genetic code.

1969 Werner Arber, Daniel Nathans and Hamilton O. Smith discover restriction enzymes.

1972 Paul Berg and others join DNA from bacteria to DNA of a virus, producing the first hybrid DNA molecule. The technique came to be known as recombinant DNA technique.

1973 Stanley Cohen and Herbert Boyer insert recombinant DNA into *E. coli* bacteria that reproduce with the inserted DNA.

1977 First human gene – for chorionic somatomammotropin – cloned.

1977 Fredrick Sanger and Walter Gilbert independently discover techniques for the rapid sequencing of the order of nucleotides in DNA molecules.

1978 Sickle-cell anaemia is diagnosed before birth of an infant by analysis of its DNA.

1980 Ananda Mohan Chakrabarty gets US patent for genetically altered bacteria that could clean up oil slicks.

1982 Human insulin produced by recombinant DNA techniques.

1983 Karry B. Mullis invents the polymerase chain reaction (PCR) that allows the multiplication of DNA fragments by billions of times in a few hours.

1985 Alec Jeffreys invents DNA fingerprinting, the use of DNA analysis to match people to biological tissues found at crime scenes.

1987 DNA paternity testing becomes commercially available.

1990 Official start of the Human Genome Project in October.

1995 First genome – that of *H. influenzae* – fully sequenced.

1996 First eukaryote genome – that of yeast – fully sequenced.

1997 First cloned mammal, Dolly, born.

2000 Draft of the complete human genome made public.

Glossary

Allele An alternate form of a gene. Usually there are two alleles for every gene, sometimes as many a three or four.

Amino acids A group of organic compounds which are the basic components of protein molecule.

Antibody A protein that binds to an invading antigen before destroying it.

Antigen A foreign substance, usually a protein, that triggers the body's immune system to produce antibodies.

Bacteriophage Any virus that infects only bacteria.

Bacterium A single-celled microorganism without a nucleus.

Base In genetics, any one of four side groups composed of carbon, oxygen, hydrogen, and nitrogen atoms attached to the sugar-phosphate backbone of the DNA molecule.

Biotechnology Technology developed for the application of biological processes to the production of materials of use in medicine and industry.

Blastocyst The first stage of development of an animal embryo as a result of division of a fertilised egg.

Central Dogma The basic belief held by molecular biologists that flow of genetic information can only occur from DNA to RNA to proteins.

Chromosome A microscopic thread-like structure made up of a long chain of DNA in which genes are located.

Clone One or more genetically identical organisms.

Codon A triplet of bases that represents one of the 20 amino acids used in the synthesis of protein molecules in the living cell.

Deoxyribose A five-carbon sugar molecule found in DNA

that contains one oxygen less than ribose.

Deoxyribonucleic acid *See* DNA

DNA Deoxyribonucleic acid, a long chain-like molecule shaped like a double-helix and composed of nucleotides – the stuff of which genes are made.

DNA chip A microarray of single-stranded DNA of known sequences attached to a silicon or glass substrate, which makes use of the complementary pairing of DNA strands to identify thousands of genes simultaneously.

DNA fingerprinting A technique in which the unique pattern formed by the nucleotide sequence of certain stretches in the DNA of an individual is matched for identification.

DNA ligase An enzyme that is used to join two strands of DNA together.

DNA polymerase An enzyme that, in combination with certain nucleotide building blocks, can produce precise replicas of short stretches of DNA in a test tube.

Dominant A term applied to the trait (allele) that is expressed regardless of the second allele.

E. coli Escherischia coli – a common bacterium found in the intestine, which has been the workhorse of research in genetics.

Enzyme An organic catalyst, usually a protein that initiates or speeds up chemical reactions in a living organism.

Eukaryote An organism made up of cells with distinct nucleus.

Exons Short stretches of DNA in a gene that code for a protein.

Gene A segment of DNA containing coded information for making a specific protein that governs transmission of a specific hereditary trait.

Gene sequencer A machine design to carry out automated biochemical tests necessary to identify genetic sequences using specially designed gene probes.

Genetic code A specific sequence of bases in DNA that guides

the synthesis of proteins by linking amino acids in the ribosome.

Genetic engineering The technique of artificially altering the genetic make-up of an organism.

Genetics The branch of biology that deals with heredity and variations in organisms, the mechanism which bring about the variations.

Genomics The study of genetic material on a large scale.

Histocompatibility antigens Carbohydrate-containing proteins found on the surfaces of almost all body cells that are involved in recognizing antigens.

Histones A group of small protein molecules found in cell nucleus that act as scaffolding for DNA in the chromosome.

Hormone A substance secreted by one type of cell that carries a signal to another type of cell.

Hybrid An offspring resulting from the crossing of parents differing in hereditary traits; the offspring often develops traits different from those of either parent.

Hydrogen bond A type of weak bond formed by electrostatic interaction between molecules.

Hydrolysis Splitting of a chemical bond by addition of a water molecule.

Immunotherapy The treatment of disease or infection by immunization.

Introns Stretches of DNA in a gene that do not code for a protein.

Jumping genes Small fragments of DNA that occasionally change position either within the same chromosome or across chromosomes.

Meiosis Cell division that produces sperm and egg cells.

Mitosis The process of cell division.

Molecular biology The study of biology at the molecular level.

Monoclonal antibodies Identical antibodies cloned from a single source and targeted for a specific antigen.

mRNA Messenger RNA, the ribonucleic acid in a living cell that becomes the working copy of a segment of a strand of DNA; serves as a mechanism for transmitting genetic instructions from the DNA in the nucleus of a cell to the protoplasm for protein synthesis.

Mutant An organism differing from its parents in one or more inheritable characteristics.

Mutation A sudden change in one or more characteristics of an organism that can be transmitted to subsequent offspring; it can be artificially caused by chemicals or ionising radiation.

Nucleic acid A complex, weakly acidic chemical compound found especially in the nucleus of all living cells; there are two forms known as RNA and DNA.

Nucleotide Building block of DNA molecule composed of a sugar, a phosphate and a nitrogen-containing organic base.

Nucleus In living cells, a dense, rounded body that contains the chromosomes.

Operator A short segment of DNA located between the promoter and the structural genes in the operon.

Operon Genetic regulatory system found in bacteria and their viruses in which genes coding for functionally related proteins are clustered along the DNA.

PCR Polymerase chain reaction, a technique that allows the multiplication of DNA fragments by billions of times in a few hours.

Plasmid A self-replicating circular DNA molecule found in bacteria and carrying two or more genes.

Pluripotent Cells capable of developing into several different cell types.

Polynucleotide A chain of several nucleotides.

Polypeptide A compound formed by linking two or more amino acids.

Promoter A region of DNA in a gene that is involved in the initiation of transcription.

Protein Any one of a large group of complex organic compounds formed by joining amino acid molecules together that are important constituents of all living cells.

Proteomics Large-scale studies of protein structure and function.

Recombinant DNA A new combination of genes spliced together on a single piece of DNA.

Recessive A term applied to a trait that is only expressed when the second allele is the same.

Repressor In bacteria, a protein which shuts down transcription of an operon by binding to the operator.

Restriction enzyme A large group of enzymes that can cut the DNA strand at specific locations.

Reverse transcriptase A kind of polymerase that can make DNA from an RNA template.

RFLP Restriction-fragment-length polymorphism, a concept employed in DNA fingerprinting.

Ribonucleic acid *see* RNA

Ribose A five-carbon sugar molecule found in RNA.

Ribosome A complex biological particle that moves along a segment of mRNA during protein synthesis, "reading" the gentic message and translating the base sequence of the mRNA into an amino acid sequence.

RNA Ribonucleic acid; a complex chemical substance found in all living cells and believed to play an important role in the synthesis of proteins and enzymes in the cell.

RNA polymerase An enzyme that, in combination with certain nucleotide building blocks, can produce precise replicas of short stretches of RNA in a test tube.

Southern blotting A technique in which DNA fragments separated by electrophoresis are transferred by "blotting" to a nitrocellulose membrane.

Totipotent Cells capable of developing into complete individuals.

Transcription The process of copying of a DNA strand, following base-pairing rules, by an RNA polymerase to produce a complementary RNA copy.

Transduction The genetic modification of a bacterium by genes of another bacterium carried by a bacteriophage.

Translation The process by which the genetic information encoded in messenger RNA directs the synthesis of specific proteins.

Transposons Genetic elements that that occasionally change position either within the same chromosome or across chromosomes.

tRNA Transfer RNA, a type of small RNA found in all cells which carries amino acids to the ribosomes for protein synthesis.

VNTR Variable number tandem repeats; sequences of base pairs found in non-coding regions of DNA, containing anywhere from 20 to 100 base pairs.

Index